

Inserm



La science pour la santé _____
_____ **From science to health**

Guide d'utilisation

Je transfère ma donnée

Comment **transférer des données en téléversement et en téléchargement** dans le Cloud Inserm pour pouvoir les exploiter dans les différentes offres du Cloud self-service

INSERM - Cloud self-service

Sommaire

1

INTRODUCTION

Enjeux des transferts de données dans le Cloud

Données HDS ou nHDS : quelles étapes suivre ?

Comment transférer mes données nHDS ?

Comprendre le fonctionnement général de la plateforme

Stockage S3 : rôle et usages

2

PARTIE 1

TELEVERSER DES DONNÉES HDS DANS LE CLOUD

3

PARTIE 2

TRAITER LES DONNÉES HDS

4

PARTIE 3

TELECHARGER DES DONNÉES HDS



INTRODUCTION

Enjeux des transferts de données dans le Cloud

Permettre l'ingestion de gros volumes

via des outils standard (clients lourds type Cyberduck, rclone, AWS CLI, etc.)

Assurer la conformité sécurité du traitement des données

le téléversement doit subir un scan antivirus

le téléchargement doit suivre un processus de validation

Données HDS ou nHDS : quelles étapes suivre ?

Pour rappel, la plateforme Cloud self-service est divisée en 2 parties :

Données nHDS

Données HDS

❏ Les étapes de ce guide concernent la plateforme pour les données **HDS**.

Allez à la page suivante pour **transférer des données non HDS**.

Comment transférer vos données non HDS ?

Vous pouvez transférer vos données via FTP ou l'outil de votre choix. Pour ce guide, prenons l'exemple du transfert de données vers une machine Linux avec WinSCP :

1 Pré-requis

1. Être en zone non HDS
2. Installer WinSCP (<https://winscp.net/eng/download.php>)

3 Connexion

1. Cliquer sur WinSCP pour lancer l'application
2. Dans la fenêtre principale, sélectionnez SFTP (SSH File Transfer Protocol)
3. Entrez l'IP ou le nom du serveur (ex. 192.168.1.10)
4. Sélectionner le port 22 (par défaut pour SFTP)
5. Rentrer un nom d'utilisateur
6. Taper votre mot de passe ou passer à l'étape 4 si vous avez une clé SSH privée
7. Cliquez sur Connexion pour vous connecter à votre serveur.

2 Connexion SSH (comme pour PuTTY) :

1. Choisir l'adresse IP ou nom du serveur
2. Taper votre nom d'utilisateur : prenom.nom
3. Taper votre mot de passe ou clé privée (.ppk ou .pem)

4 Connexion avec une clé SSH privée

1. Cliquer sur "Avancé..."
2. Allez dans "Authentification" (sous SSH)
3. Dans "Fichier de clé privée", sélectionnez votre clé au format .ppk
4. Si vous avez une clé .pem, convertissez-la d'abord avec PuTTYgen
5. Cliquez sur OK pour revenir à la fenêtre principale.
6. Cliquez sur Connexion pour vous connecter à votre serveur.

5 Transfert de fichiers entre votre ordinateur et le serveur Linux

Glisser-déposer les fichiers à transférer ou utiliser les boutons de transfert selon votre convenance

Comprendre le fonctionnement général de la plateforme Cloud self-service en environnement HDS

Votre environnement repose sur **4 buckets distincts**, chacun ayant un rôle précis :

Bucket	Accès Internet	Droits	Rôle
Bucket d'entrée	Oui (via F5 APM)	Écriture seule	Déposer vos données et effectuer un scan antivirus
Bucket de travail	Non	Lecture / écriture	Traiter vos données (HPC, outils cloud)
Bucket de sortie temporaire	Non	Lecture / écriture	Permettre le processus de validation
Bucket de sortie permanent	Oui (via F5 APM)	Lecture seule	Télécharger vos résultats

📄 🔒 L'usage du VPN remonte.hds.inserm.fr pour accéder aux buckets **est optionnel**. Il n'est pas obligatoire mais peut vous permettre de sécuriser une bande-passante plus importante.

Stockage S3 : rôle et usage

Le stockage S3 n'est pas destiné aux mêmes usages que le stockage capacitif sur SSD.

Le **S3** est un stockage objet adapté à l'archivage, au dépôt de données et aux usages ne nécessitant pas de hautes performances. Il peut aussi être utilisé comme **espace transitoire**, notamment pour transférer des données :

- vers le **HPC** ;
- vers le **stockage capacitif**, pour une exploitation ultérieure à plus haute performance.

❏ Le **stockage capacitif sur SSD** est, quant à lui, à privilégier pour les usages de travail nécessitant de meilleures performances d'accès aux données.

❏ Le stockage S3 est soumis à des **limites de capacité**, à la fois **individuelles** et **globales** à l'infrastructure. Les utilisateurs sont invités à veiller à l'adéquation entre leur besoin et le type de stockage mobilisé.

TELEVERSER DES DONNÉES HDS DANS LE CLOUD INSERM

Étape 1 – Récupérer les informations du bucket d'entrée

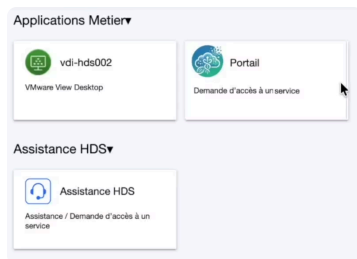
01

Se connecter

Se rendre sur la page d'accueil du Cloud HDS

[:portail.hds.inserm.fr](https://portail.hds.inserm.fr)

Cliquer sur le portail de commandes



02

Créer

Depuis la page d'accueil, cliquer sur le projet dans lequel vous souhaitez téléverser des données

Cliquer sur "Mes services"

Cliquer sur "Import de données"

Cliquer sur "1. Activer le service d'import"

Vous avez ainsi l'URL de votre bucket.

03

Récupérer

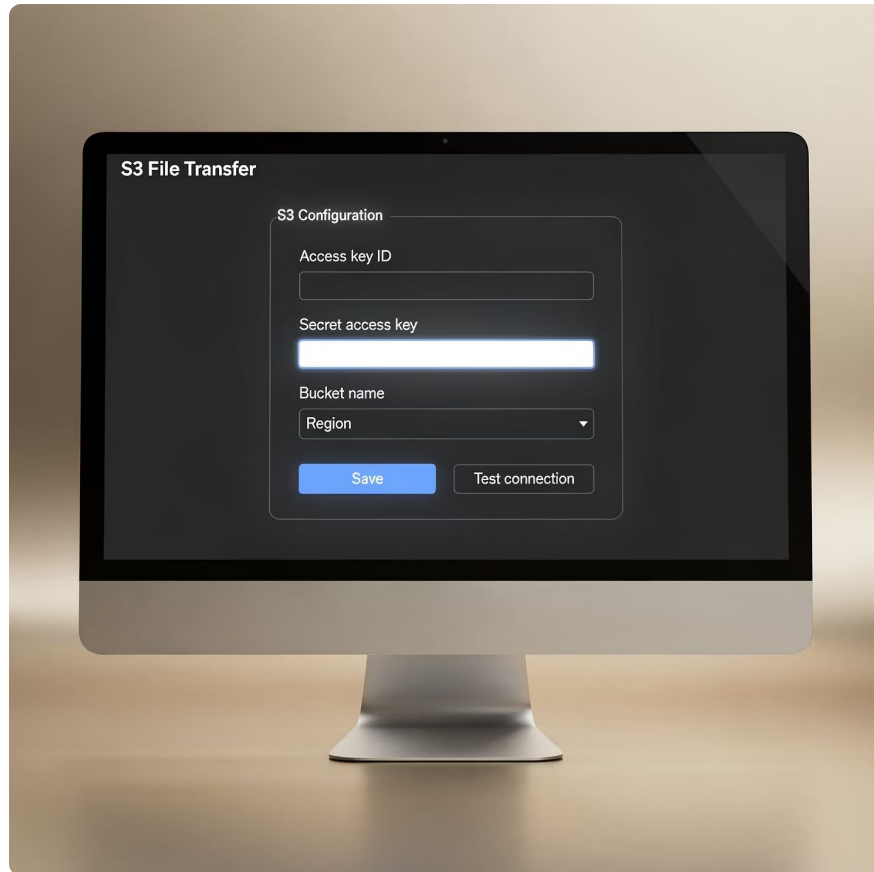
Cliquer sur "2. Créer/ Régénérer les clés S3" pour récupérer les clés de votre bucket sur la page de votre projet :

- Access Key (AK)
- Secret Key (SK)

ⓘ ⚠ Si vous ne vous souvenez pas de vos identifiants, vous pouvez les réinitialiser en cliquant sur "2. Créer. Régénérer les clés S3".

⚠ Attention, les identifiants de l'ensemble de vos buckets seront réinitialisés.

Étape 2 – Configurer votre outil de transfert



1. **Démarrer** le client de transfert de données compatible S3 de votre choix (Cyberduck, FileZilla, rclone, AWS CLI, etc.) sur votre poste.
2. **Saisir** les informations d'identification suivantes :
 - L'URL du bucket
 - Votre clé d'accès (AK)
 - Votre clé secrète (SK)
3. **Valider** la configuration.

Étape 3 – Déposer les données

1. Téléverser

Déposer les données que vous souhaitez téléverser dans votre outil.



3. Valider

Cliquer sur "3. Terminer et lancer le scan antivirus".



2. Vérifier

Vérifier que le dépôt est terminé sur votre client lourd.

Étape 4 – Scanner et téléverser vos données sur le bucket de travail

☐ Vous **n'avez rien à faire** : le scan antivirus et le transfert vers le bucket de travail se fait de manière automatique.

 Vous pouvez **consulter le statut du scan** depuis le portail en vous rendant sur le projet concerné et en cliquant sur "Mes services".

➔ **Une fois le scan validé :**

Vos données sont copiées vers votre bucket de travail.

☐  Une fois les données transférées dans le bucket de travail, le bucket d'entrée est vidé automatiquement.

TRAITER LES DONNÉES HDS

Étape 5 – Accéder au bucket de travail

Accéder et exploiter

- Pour exploiter les données sur le HPC, voir le guide "HPC bucket S3"
- Pour exploiter les données avec l'infrastructure, voir le guide "offre infrastructure"

 Le bucket de travail n'est pas accessible depuis Internet. Les traitements s'effectuent **exclusivement depuis l'infrastructure interne.**

TELECHARGER DES DONNÉES HDS

⚠ En environnement HDS, la sortie est soumise à validation.

Étape 6 – Préparer la sortie



1. Créer

- Depuis la page d'accueil, cliquer sur le projet dans lequel vous souhaitez télécharger des données
- Cliquer sur "Mes services"
- Cliquer sur "Export de données"
- Cliquer sur "1. Activer le seuil de service d'export"



2. Récupérer

- Cliquer sur "2. Créer/ Régénérer les clés S3"
- Récupérer les identifiants du bucket de sortie temporaire (URL, AK, SK)



3. Déposer

- Déposer les fichiers à exporter dans le bucket de sortie temporaire depuis vos outils Cloud Inerm

Étape 7 – Lancer la validation

1. Vérifier

Consulter vos mails pour vérifier que le dépôt est terminé.
L'émetteur du mail est le suivant : noreply@inserm.fr



3. Informer

Un email est envoyé automatiquement au valideur de la donnée défini sur votre projet. Nous vous recommandons toutefois de le contacter par vos propres moyens pour s'assurer de sa validation dans les temps.



2. Valider

Cliquer sur "3. Terminer l'export" pour valider la fin du dépôt sur le portail de commandes



📄 Une fois la demande de validation acceptée, les fichiers sont analysés par l'antivirus avant mise à disposition puis les données sont automatiquement copiées vers le bucket de sortie permanent.

🗑️ Le bucket de sortie temporaire est vidé automatiquement une fois la donnée transférée vers le bucket de sortie permanent.

Étape 8 – Télécharger les résultats

1

Récupérer

Récupérer les identifiants du bucket de sortie permanent en cliquant sur le projet concerné (URL, AK, SK).

2

Configurer

- Démarrer le client de transfert de données compatible S3 de votre choix (Cyberduck, FileZilla, rclone, AWS CLI, etc.) sur votre poste.
- Saisir les informations de votre bucket (URL, AK, SK)
- Valider la configuration

3

Télécharger

Télécharger vos résultats en lecture seule.

📋 N'oubliez pas de noter les identifiants de votre bucket de sortie permanent pour les réutiliser.

⚠️ Seules les méthodes GET/ HEAD sont autorisées.