



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

Guide d'utilisation

VM de staging HPC

Ce guide vous permettra de comprendre les cas d'usage d'une VM de staging.

INSERM - Cloud self-service

Agenda

- 1 Définition et bénéfices d'une VM de staging
- 2 Ce que la VM de staging vous permet de faire
- 3 Ce que la VM de staging ne vous permet pas de faire
- 4 Exemple d'un workflow typique
- 5 Pour aller plus loin

Définition et bénéfices d'une VM de staging

À quoi sert une VM de staging ?

La VM de staging est un **espace de préparation** mis à votre disposition avant de lancer vos calculs sur la plateforme HPC pour :

- Tester vos outils
- Préparer vos environnements
- Construire vos conteneurs

Quels sont les bénéfices de la VM de staging ?

- Débuguer vos pipelines avant exécution sur le HPC
- Itérer rapidement sans passer par le scheduler Slurm
- Garantir la reproductibilité des environnements (via apptainer)

Ce que la VM de staging vous permet de faire

Préparer vos outils

- Installer des dépendances avec **conda**, **pip** ou **R** depuis le serveur Nexus de l'Inserm
- Tester votre environnement logiciel

Construire des conteneurs

- Créer des images **Singularity / Apptainer**
 - Convertir une image **Docker** → **.sif**
- 👉 Utile car Docker n'est pas disponible sur le HPC

Utiliser des outils spécifiques

- Utiliser Docker (uniquement sur la VM de staging)
- Utiliser des outils de compilation (gcc, make...)
- Installer des paquets spécifiques RHEL 9 via les dépôts Satellite Inserm

Préparer vos workflows

Utiliser **Nextflow** pour organiser vos pipelines

- 📄 Nextflow peut aussi piloter le HPC depuis un noeud de login.

Gérer vos données

- Envoyer vos données et images vers la plateforme HPC via un **bucket S3**
- Préparer vos fichiers avant exécution sur le HPC

Ce que la VM de staging ne vous permet pas de faire

Calculs intensifs

Les ressources de la VM de staging mises à disposition ne sont pas prévues pour lancer des calculs intensifs.

Accès direct HPC

Les données de votre compte HPC se sont pas accessibles depuis la VM de staging.

Accès direct SLURM

Il n'est pas possible d'accéder directement au scheduler SLURM, car les données de votre compte HPC se sont pas accessibles depuis la VM de staging.

Exemple d'un workflow typique

Vous trouverez ci-après un exemple de workflow typique :

0. Commander et se connecter à la VM de staging

1. Préparez votre environnement sur la VM


2. Créez une image .sif

3. Transférez vos données et vos images SIF vers un bucket S3

4. Connectez vous au HPC (nœud de login)

5. Importez les données depuis le bucket S3 vers votre espace HPC

6. Lancez votre calcul avec SLURM depuis votre espace HPC

 Liste des outils préinstallés :

- Cortex
- Microsoft Office 2024
- Google Chrome
- 7-Zip
- NotePad++
- Adobe Acrobat Reader
- MobaXTerm
- WinSCP
- DBeaver
- Microsoft Visual Studio 2022 + VS Code + Plugins Python / R / Java / Julia
- RStudio

Pour aller plus loin

1

Passer de Docker à Apptainer

INSERM Cloud - Tutoriel Slurm (chapitre "De Docker à Apptainer")

2

Accéder au S3

INSERM Cloud - Tutoriel HPC exploitation des données S3

3

Utiliser Conda

INSERM Cloud - Tutoriel HPC utilisation de Conda

4

Installer les biocontainers

INSERM Cloud - Tutoriel HPC installation de Biocontainers