

Module HAT918T
Outils de modélisation hydro-morphodynamique littorale et portuaire
**Installation et lancement de DELFT3D WAVES
sur cluster, découverte du format netCDF**

Session 04-05

Louis Saisset

Ce TD est la suite de la prise en main des outils de modélisation hydro-morphodynamique littorale et portuaire. Pour prendre en main SWAN en 2D-horizontal, nous utiliserons le module WAVES de la suite logiciel DELFT3D.

Introduction

Connectez vous sur le cluster via `ssh` comme lors du tp précédent. Dans le répertoire de travail, vous trouverez le dossier `TP_Hydro_D3D`. Avant de vous lancer dans le reste du TP, copiez ce dossier dans votre répertoire personnel.

```
1 \ $ cp -r ~/work_f_gcl/TP_Hydro_D3D/ ~/
```

Vérifiez le contenu du dossier `TP_Hydro_D3D`. Vous devriez y trouver :

- Un dossier **Installation_Rapide** contenant 1 dossier de chemins pour module, un fichier de commande et 2 fichiers cachés permettant de faciliter l'utilisation du cluster (pour voir les fichiers cachés, utilisez la commande `ls -la`).
- Un dossier **Simulation** contenant une configuration complète prête à être lancée.
- Un dossier de **Resultats** à post-traiter sous python correspondant aux résultats de la simulation fournie.
- Un dossier **Mesures** contenant des prises de mesures et des localisations (3 fichiers au total).
- Un fichier contenant des informations **Utiles** (fonctions python, sites d'aide, etc).

1 Premiers pas avec D3D

1.1 Installation Rapide

L'objectif de cette installation rapide est de vous épargner la fastidieuse installation de Delft3D. L'installation complète, si jamais vous souhaitez la réaliser sur votre machine personnelle, passerait par l'utilisation du logiciel PAGURE¹ comme lors du TP précédent. Pour échapper aux longues heures d'installation de ce logiciel (beaucoup plus lourd que SWAN), une version de Delft3D a déjà été installée sur le groupe de travail commun. Il ne vous reste qu'à récupérer et positionner les modules créés par PAGURE dans votre répertoire personnel.

1.1.1 Fonctionnement de Module et principe de l'installation rapide

Pour pouvoir charger et décharger différentes configurations de compilation dans un terminal, le logiciel `module`² est couramment utilisé. Il permet, via un ajout (ou une suppression) de chemins, de donner l'accès à des fichiers binaires de votre choix pour votre environnement actuel. Pour cela, plusieurs commandes sont à connaître :

```
1  \ $ module purge # vide les modules chargés dans cet environnement
2  \ $ module list # affiche l'ensemble des modules actuellement chargés dans cet environnement
3  No Modulefiles Currently Loaded.
4  # Pour le moment il n'y a rien puisqu'on vient de le vider...
5  \ $ module avail # affiche les modules qui PEUVENT être chargés dans cet environnement
6  ----- /usr/share/Modules/modulefiles -----
7  dot          module-git  module-info  modules      null          use.own
8  # ... plein de trucs ...
9  # Tous ces modules sont classés en fonction d'une localisation,
10 # indiquée par le chemin d'un dossier.
11 \ $ module load cv-standard # charge cv-standard en mémoire pour cet environnement
12 \ $ module list
13 Currently Loaded Modulefiles:
14   1) cv-standard
15 # On voit le module qu'on vient de charger
16 \ $ module load use.own # charge le dossier /privatemodules en mémoire pour cet environnement
17 \ $ module list
18 Currently Loaded Modulefiles:
19   1) cv-standard  2) use.own
20 # Le deuxième module vient s'ajouter à la suite du premier
21 \ $ module avail
22 # Même chose qu'au avail précédent avec 2 onglets supplémentaires : cv-standard et privatemodules
```

Vous remarquerez que `module load use.own` donne l'accès aux modules contenus dans le dossier `/home/e_gcl-XX/privatemodules`. Tous les logiciels installés par PAGURE sur votre session personnelle correspondent à des ajouts de modules dans CE dossier uniquement (assimilable à des ajouts de chemins disponibles). Pour que `module` vous donne l'accès au logiciel souhaité, il faut donc que votre dossier `privatemodules` contienne l'entièreté des chemins pointant vers le lieu d'installation du logiciel.

1. <https://github.com/fretif/pagure>
2. <https://modules.readthedocs.io/en/latest/>

EXEMPLE : Lors du TP précédent³, vous avez été amenés à installer SWAN sur votre session personnelle. En utilisant PAGURE, vous avez enregistré, d’une part le code de SWAN dans le dossier `/home/e_gcl-XX/softs`, et d’autre part vous avez enregistré les modules permettant d’accéder à ce code dans votre dossier `/home/e_gcl-XX/privatemodules`. Vous aviez alors accès au chargement de SWAN en mémoire grâce à la commande `module load swan`. Renommez le dossier `/home/e_gcl-XX/privatemodules` créé lors de l’installation de SWAN en `/home/e_gcl-XX/privatemodules-swan`.

```
1  \ $ mv /home/e_gcl-XX/privatemodules /home/e_gcl-XX/privatemodules-swan
2  \ $ module load use.own
3  \ $ module avail
4  # Aucun accès à swan n'est disponible car module ne voit plus le dossier /privatemodules.
5  # L'installation de SWAN n'a pourtant pas été supprimée ! Module a juste perdu son accès !
6  # Pour que module récupère cet accès, il suffit de rendre au dossier son nom d'origine.
```

Si vous aviez réalisé une installation complète de Delft3D sur votre session, PAGURE se serait chargé de construire des chemins adaptés directement dans votre dossier `privatemodules`. Pour éviter d’effacer les chemins utiles pour les autres logiciels (chose qu’aurait fait PAGURE à votre place), vous allez plutôt conserver vos ”chemins SWAN” dans un dossier `privatemodules-swan` et copier le dossier `privatemodules-delft` vers votre répertoire personnel. Il contiendra les chemins vers l’installation commune de Delft3D. Ainsi, `module` permettra l’accès aux chemins contenus le dossier que vous choisirez de renommer `privatemodules`. Vous évitez ainsi d’écraser les anciens chemins (stockés dans un dossier à part) tout en ajoutant les chemins nécessaires au lancement de Delft3D (stockés dans le dossier visible par `module`).

```
1  \ $ cp -r ~/TP_Hydro_D3D/Installation_Rapide/privatemodules-delft ~/
2  \ $ mv ~/privatemodules-delft ~/privatemodules
```

A ce stade, il vous est déjà possible d’accéder à Delft3D. Videz la mémoire de `module` puis vérifiez qu’il n’y a plus aucun module de chargé (ligne 1 à 3). Chargez ensuite les dossiers `cv-standard` et `use.own` qui donnent accès respectivement aux logiciels communs de compilation et aux logiciels du répertoire personnel `privatemodules` (lignes 4 à 7). Puis vérifiez que `delft3d` est bien dans la liste des logiciels accessibles par `module` (ligne 8 et plus) dans l’onglet `/home/e_gcl-XX/privatemodules` :

```
1  \ $ module purge
2  \ $ module list
3  No Modulefiles Currently Loaded.
4  \ $ module load cv-standard use.own
5  \ $ module list
6  Currently Loaded Modulefiles:
7     1) cv-standard   2) use.own
8  \ $ module avail
9  # beaucoup de modules
10 ----- /home/e_gcl-19/privatemodules -----
11 delft3d/mpich321/icc17/6.03
12 # beaucoup de modules
```

3. voir TP1 partie 2 : Déployer SWAN sur le cluster

1.1.2 Facilitation d'accès

A chaque ouverture d'un terminal, la liste des logiciels accessibles par *module* se décharge. Pour pouvoir lancer des simulations, il nous faudra donc charger `cv-standard` et `use.own` à chaque ouverture d'un terminal (ou chaque nouvelle connexion au cluster).

Pour faciliter le lancement de Delft3D sur le cluster, 3 autres fichiers vous ont été fourni : un fichier exécutable `submitter_dwave_parallel.cmd` et deux fichiers cachés `.bash_profile` et `.bashrc`. Les fichiers cachés ne sont pas visibles par défaut mais peuvent l'être avec la commande `ls -la` :

```
1  \$$ cd ~/TP_Hydro_D3D/Installation_Rapide/
2  \$$ ls
3  privatemodules-delft submitter_dwave_parallel.cmd
4  \$$ ls -la
5  drwxr-sr-x  3 e_gcl-19 f_gcl  122  6 oct.  16:23 .
6  drwxr-sr-x  6 e_gcl-19 f_gcl  114  6 oct.  16:21 ..
7  -rw-r--r--  1 e_gcl-19 f_gcl  194  6 oct.  15:54 .bash_profile
8  -rw-r--r--  1 e_gcl-19 f_gcl  168  6 oct.  16:22 .bashrc
9  drwxr-xr-x 11 e_gcl-19 f_gcl 4096  5 oct.  16:14 privatemodules-delft
10 -rwxr-xr-x  1 e_gcl-19 f_gcl  950  5 oct.  17:56 submitter_dwave_parallel.cmd
```

Copiez le fichier exécutable dans un dossier `TOOLS` à la racine de votre espace personnel et copiez les fichiers cachés directement à la racine.

```
1  \$$ mkdir ~/TOOLS # crée le dossier /TOOLS à la racine
2  \$$ cp ~/TP_Hydro_D3D/Installation_Rapide/submitter_dwave_parallel.cmd ~/TOOLS/
3  \$$ cp ~/TP_Hydro_D3D/Installation_Rapide/.bashrc ~/
4  \$$ cp ~/TP_Hydro_D3D/Installation_Rapide/.bash_profile ~/
5  \$$ source .bashrc
6  # Nécessaire pour que l'environnement actuel prenne en compte les changements que vous venez
7  # d'effectuer.
```

En lisant le fichier `.bashrc` on voit qu'il charge les dossiers `cv-standard` et `use.own` puis qu'il ajoute un accès rapide au dossier `TOOLS`. L'autre fichier permet le lancement automatique de `.bashrc` lors de l'ouverture de chaque nouvel environnement.

Pour comprendre l'utilité de cette manipulation, déconnectez vous puis reconnectez vous au cluster. Vous verrez alors (`module avail`) que `use.own` et `cv-standard` sont immédiatement disponibles. De même, où que vous vous positionniez dans l'arborescence, la commande `submitter_dwave_parallel.cmd` doit être disponible dans l'auto complétion de votre terminal (dans le terminal, tapez `sub` puis complétez avec la touche `tab`).

1.2 Premiers lancements

1.2.1 Delft3D

EXPLICATIONS AU TABLEAU

1.2.2 Utiliser un Cluster de calcul

Si cela n'est pas déjà fait, connectez vous sur le cluster via `ssh`.

Quand vous interagissez avec le cluster, par exemple quand vous déplacez ou ouvrez des fichiers, vous êtes connecté sur les machines `muse-login01` ou `muse-login02`. Vous utilisez par conséquent les ressources de ces 2 noeuds de connexion lors de toutes vos actions. Ces machines sont les seules interfaces de connexion et sont partagées par tous les utilisateurs du cluster. Toute action réalisée sur ces noeuds impacte donc vos camarades, mais aussi l'ensemble des équipes de recherche connectées en même temps que vous sur ces noeuds. Si vous utilisez la commande `htop` vous pouvez d'ailleurs voir l'ensemble des actions en cours sur ce noeud ainsi que les capacités disponibles. Si vous avez de la chance, il vous sera possible de prendre en flagrant délit les utilisateurs saturant l'espace commun.

Dans le TP précédent, les calculs lancés sur SWAN étaient très légers et très rapides. Dans cette optique, les lancements ont tous été réalisés directement sur les noeuds de connexion. Durant la suite de ce TP, cette démarche sera à proscrire pour éviter de saturer complètement le cluster. Les noeuds de connexion ne sont en aucun cas des noeuds de calcul !

Pour utiliser les noeuds de calcul du cluster, il faut impérativement passer par le gestionnaire de tâches **SLURM**. C'est un gestionnaire de ressources permettant de réserver puis d'assigner des ressources de calcul à un travail. Pour bien comprendre la différence, connectez vous au cluster sur deux terminaux différents.

```
1  # TERMINAL 1 :
2  \$ cd ~/TP_Hydro_D3D/Simulation/ # déplacez vous vers le dossier Simulation
3  \$ submitter_dwave_parallel.cmd  # lancez le calcul directement depuis le dossier
4  # vous voyez quelque chose se lancer...
5
6  # TERMINAL 2 :
7  \$ htop # lancez l'interface et regardez l'état de votre noeud de connexion
8  # Vous devriez constater que vous êtes en train de saturer un noeud !
9  # Arrêtez moi ce calcul que je ne saurais voir !!!
10 # Vous bloquez tout le monde !!!
11 # (tapez ctrl+c dans le premier terminal et tout ira mieux)
12
13 # TERMINAL 1 :
14 \$ cd ~/TP_Hydro_D3D/Simulation/ # retournez dans le même dossier
15 \$ sbatch submitter_dwave_parallel.cmd # demandez à SLURM de réserver de la place pour vous
16 Submitted batch job XXXXXXXX
17 \$ watch squeue
18      JOBID PARTITION      NAME      USER ST TIME  NODES NODELIST(REASON)
19      XXXXXXX      gladys test_DWA e_gcl-XX R 1:19      1 muse036
20 # vous avez demandé a réaliser le calcul et SLURM l'a lancé sur le noeud muse036
```

Vous avez maintenant lancé votre première simulation delft3d sur cluster.

2 Post-Traitement des données et découverte du format netCDF

L'objectif de ce TP est de vous faire manipuler delft3d à la fois au niveau de son lancement, mais aussi au niveau du traitement des sorties du modèle.

Commencez par installer les librairies python `pandas`, `numpy`, `scipy`, `pyproj`, `netCDF4` et `matplotlib` sur votre machine car elles vous seront nécessaires par la suite, puis téléchargez sur votre machine le contenu des dossiers `Resultat` et `Mesures` avec la commande `scp` comme lors du TP précédent.

En vous aidant des commandes et des liens indiqués dans le fichier "Utile" réalisez les étapes suivantes :

1. Charger les données netCDF `wavm-maupiti.nc` dans une variable.
2. Explorer l'architecture netCDF de cette variable pour en extraire les valeurs des profondeurs.
3. Afficher la carte des profondeurs en fonction de leurs coordonnées. Attention, les points en $x=0$ et $y=0$ ne doivent pas être affichés.
4. Extraire les noms et les localisations des appareils de mesure du fichier contenant en colonne : nom, longitude et latitude.
5. Afficher la localisation des appareils de mesure sur cette carte. Leurs coordonnées doivent être mises en convention EPSG32705.
6. Afficher une carte de la hauteur significative pour l'évènement modélisé.
7. Extraire les mesures de Hs pour le profil SW et les afficher en fonction de leur date.
8. Comparer les mesures terrain avec le résultat de la modélisation pour chacune des 6 sondes.
9. Représenter la coupe complète du profil Sud-Ouest (interpolation du profil de fond et des résultats de Hs depuis la sonde S4 jusqu'à la sonde OSS6)

BON COURAGE!

INDICES :

1. Lire le fichier "Utile"
2. La grille des profondeurs est dans une variable nommée "depth" quelque part dans wavm-maupiti.nc
3. Extraire les variables x, y et depth puis créer un masque de points actifs (dont x et y sont tous deux non nuls).
4. Re-lire le fichier "Utile".
5. Les coordonnées du système géographique correspondent à l'EPSG4326.
6. la variable s'appelle "hsign" dans les données netCDF (même procédé que pour les profondeurs)
7. Re-re-lire le fichier "Utile" dans le détail.
8. La date exacte de l'évènement simulé est indiquée dans un fichier encore inutilisé. Lisez le fichier .sp2
9. Cette question contient trois étapes. Créer un liste de coordonnées correspondant au profil souhaité (par exemple avec une régression linéaire ou une formule de droite affine). Interpoler les valeurs modélisées en chacune de ces coordonnées. Réaliser les affichages voulus.