



Utilisation de Mantis

Traitement massif de données GNSS en double-différence

22 juin 2021

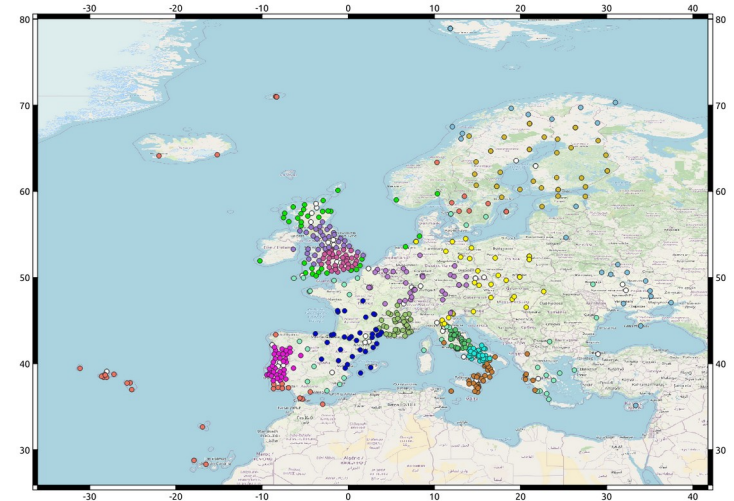
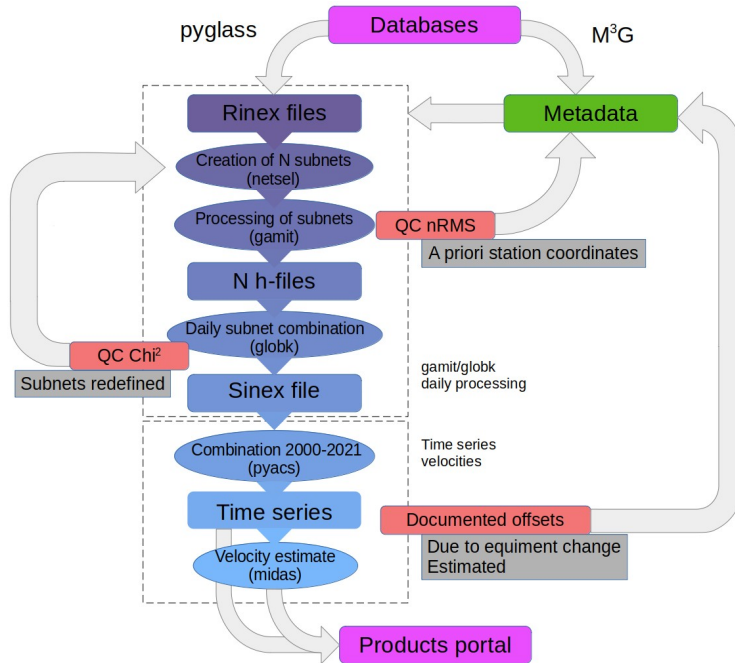


Traitement massif de données GNSS en double-différence : besoin de la grille de calcul

GNSS = Global Navigation Satellite System

Production d'une position quotidienne pour ~600 stations GNSS permanentes en Europe

Un service de l'UGA/ISTerre pour l'infrastructure EPOS (European Plate Observing System)



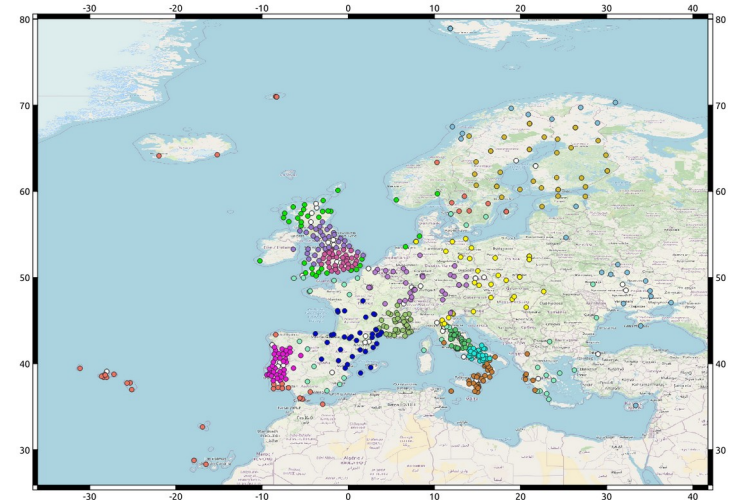
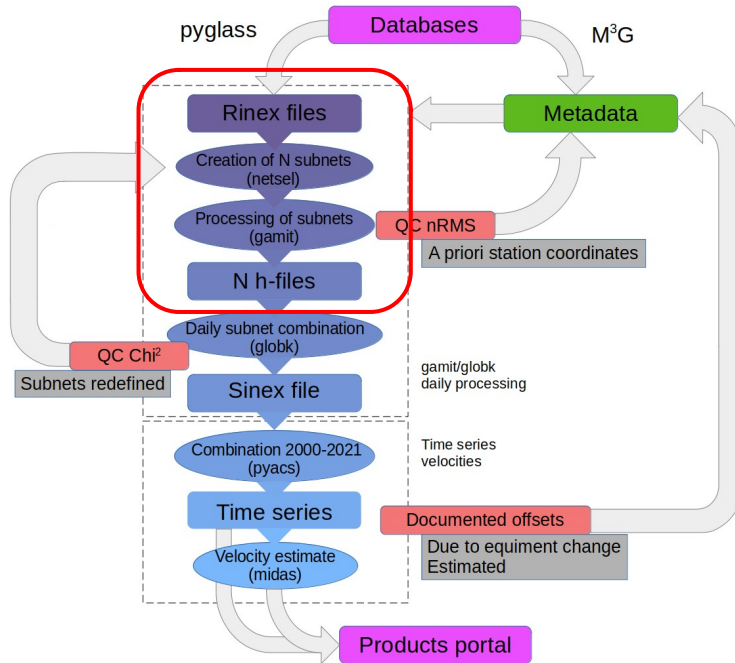
Map of 572 stations active on Aug 10th 2020

Traitement massif de données GNSS en double-différence : besoin de la grille de calcul

GNSS = Global Navigation Satellite System

Production d'une position quotidienne pour ~600 stations GNSS permanentes en Europe

Un service de l'UGA/ISTerre pour l'infrastructure EPOS (European Plate Observing System)



Map of 572 stations active on Aug 10th 2020

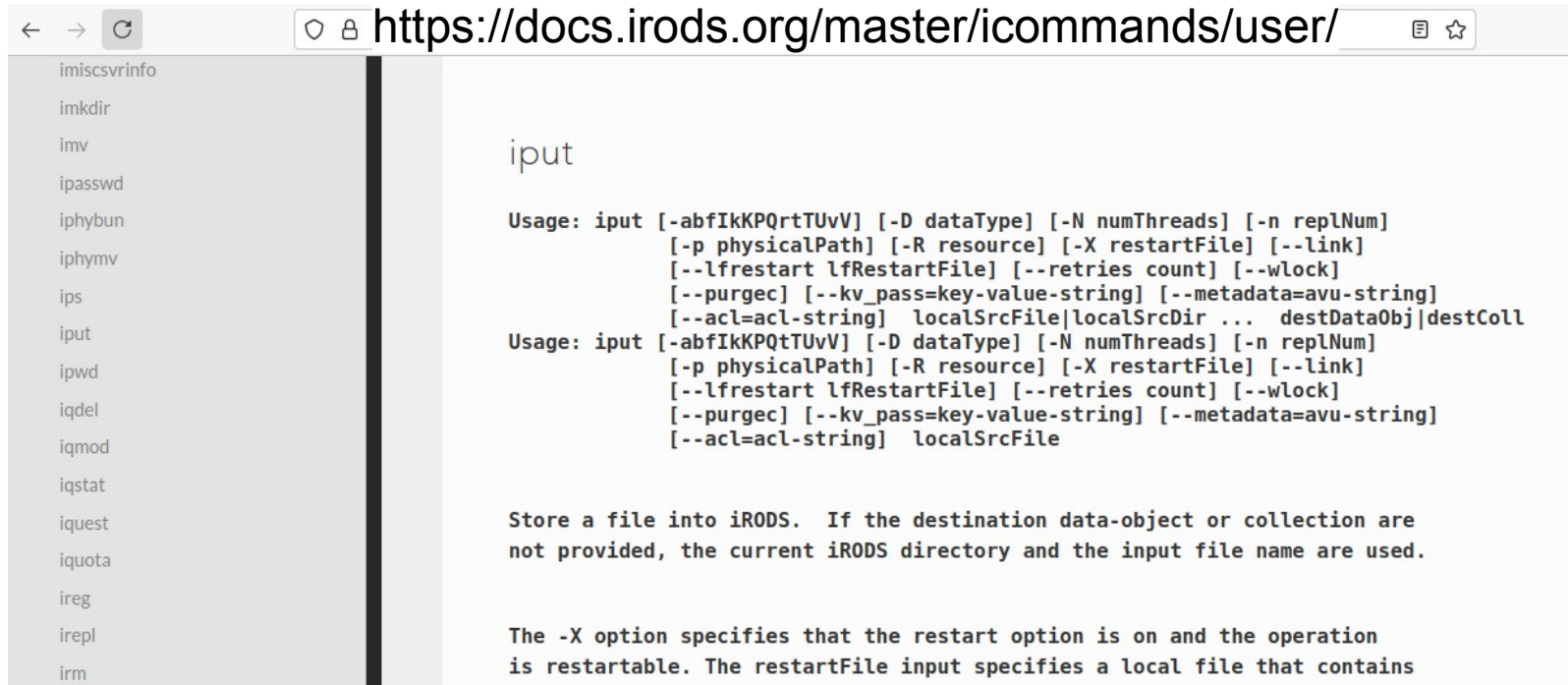
21 ans de données quotidiennes à traiter: $21 \times 365 \times 15 \sim 80000$ calculs de 30 min

Qu'est-ce que Mantis ?

- Stockage accessible depuis les différents nœuds de la grille de calcul cigri
- Mis en place dans ce but: fonctionnement avec cigri
- Utilise iRODS, qui fait l'interface avec de multiples et divers types et lieux de stockage
- Opérations de lecture / copie / transfert etc avec des « icommands » bien documentées et inspirées des commandes linux (ls, cp, rsync => ils icp/iget irsync)

Qu'est-ce que Mantis ?

- Stockage accessible depuis les différents nœuds de la grille de calcul cigri
- Mis en place dans ce but: fonctionnement avec cigri
- Utilise iRODS, qui fait l'interface avec de multiples et divers types et lieux de stockage
- Opérations de lecture / copie / transfert etc avec des « icommands » bien documentées et inspirées des commandes linux (ls, cp, rsync => ils icp/iget irsync)



← → ↻ <https://docs.irods.org/master/icommands/user/> 📄 ☆

- imiscsvrinfo
- imkdir
- imv
- ipasswd
- iphybun
- iphymv
- ips
- iput**
- ipwd
- iqdel
- iqmod
- iqstat
- iquest
- iquota
- ireg
- irepl
- irm

iput

```
Usage: iput [-abfIkKPQrtTUvV] [-D dataType] [-N numThreads] [-n replNum]
          [-p physicalPath] [-R resource] [-X restartFile] [--link]
          [--lfrestart lfRestartFile] [--retries count] [--wlock]
          [--purgec] [--kv_pass=key-value-string] [--metadata=avu-string]
          [--acl=acl-string] localSrcFile|localSrcDir ... destDataObj|destColl
Usage: iput [-abfIkKPQrtTUvV] [-D dataType] [-N numThreads] [-n replNum]
          [-p physicalPath] [-R resource] [-X restartFile] [--link]
          [--lfrestart lfRestartFile] [--retries count] [--wlock]
          [--purgec] [--kv_pass=key-value-string] [--metadata=avu-string]
          [--acl=acl-string] localSrcFile
```

Store a file into iRODS. If the destination data-object or collection are not provided, the current iRODS directory and the input file name are used.

The -X option specifies that the restart option is on and the operation is restartable. The restartFile input specifies a local file that contains

Qu'est-ce que Mantis ?

- Stockage accessible depuis les différents nœuds de la grille de calcul cigri
- Mis en place dans ce but: fonctionnement avec cigri
- Utilise iRODS, qui fait l'interface avec de multiples et divers types et lieux de stockage
- Opérations de lecture / copie / transfert etc avec des « icommands » bien documentées et inspirées des commandes linux (ls, cp, rsync => ils icp/iget irsync)

Quelques limitations...

- Pas un lieu de stockage, à voir comme un intermédiaire summer ↔ cigri
- Volume limité (quota 1TB mais peut être relevé)
- Des problèmes de corruption (?) sur mantis1 (ancienne génération) pour certains fichiers, mais on s'en sort avec le support gricad
- Quelques limitations associées aux icommands (pas trouvé d'équivalent à "rsync --delete" par exemple)
 - ...et "du -sh /myDir/mySubdir"
 - => iquest "select sum(DATA_SIZE) where COLL_NAME = '/myDir/mySubdir'"
- iinit régulier (2 semaines?) sinon erreur...

Quand utiliser Mantis ?

- Pas indispensable si vous travaillez exclusivement sur des clusters sur lesquels summer est monté (différence de performances?)
 - A court terme on pourrait se passer de mantis (actuellement calculs sur luke seulement, et summer est monté sur luke)
 - ... mais nombre de calculs par jour amené à augmenter donc on utilise mantis en vue de cette hausse pour pouvoir utiliser d'autres clusters comme froggy...

Au quotidien:

- Pré-traitement sur luke, création des jdl, param files, ...
 - ...puis mise à jour de tous les fichiers input au calcul sur mantis
 - => série de irsync et icp, pour l'instant à la main, mais à terme dans un script bash
 - ...cf démo vidéo
- Soumission cigri depuis killeen
- Récupération des fichiers de sortie sur mantis (résultats, logs, ...)
 - => iput depuis un script bash

```
mkdir -p /mantis/gnss/epos/results/${year}/${doy}
```

```
fatal_file="${results_dir}/${year}/${doy}/GAMIT_${doy}_${subnet}.fatal"  
ils $fatal_file >> /dev/null 2>&1 && irm $fatal_file
```

```
iput -f ${local_dir}/${working_dir}/${analysis}/sh_gamit_${doy}_${subnet}.summary  
      ${results_dir}/${year}/${doy}  
iput -f ${local_dir}/${working_dir}/${analysis}/GAMIT.fatal  
      $fatal_file
```

- Submission eigt depuis kinech
- Récupération des fichiers de sortie sur mantis (résultats, logs, ...)
=> iput depuis un script bash