



CMIP6 Tech

- multi-modèle
- publications
- ressources
- production
- analyse

13 Septembre 2018

Accès à ESPRI (Ciclad – ClimServ) et analyse multi-modèles

- *Financement partiel identifié (Clim-ERI)*
- *Test de la procédure d'ouverture de compte à relancer : ...*
- *ESPRI permet(tra) de voir les « file systems »*
 - *de Ganymède (5 Po) / de notre datanode ESGF / des données multi-modèles CMIP5 déjà gérées à l'IPSL*
- *La pratique IPSL est que les chercheurs émettent une sorte de 'ticket' pour demander que telle donnée CMIP soit répliquée sur l'espace commun (ici Ganymède) ; nous aurons a priori accès à cette procédure*
- *Volume du scratch prévu : 200 Go/user*
- *Le cas le moins confortable est celui où l'on veut analyser une simu CNRM-CM non-CMIP6 en comparaison de données multi-modèles : il faut alors copier la donnée maison vers ESPRI*
- **Beta-testeurs ESPRI identifiés : David St-M, ..**
- **Journée Clim-Eri information/formation/TP**
 - ▶ *L'objectif général est de faire en sorte que les personnes qui participent à cette réunion aient les informations nécessaires pour commencer à faire des analyse de données mutli-modèles sur la base de donnée CMIP6.*
 - ▶ *Le planning général pourrait être le suivant:*
 - ▶ *- matin, présentation générale de la base de données, des outils, etc...*
 - ▶ *- l'aprem, TP autour de quelques questions particulière, étude de cas, etc... Chacun viendrait avec son portable (on pourra aussi en prêter un certain nombre, sous Linux), avec son login sur ciclad, et la réunion aura lieu dans une salle avec wifi pour pouvoir se connecter*
- **5 ou 19 novembre**
- **Paris - Visio peut-être possible**
- **Ouverture de compte pour tous à cette date : pas acquise**

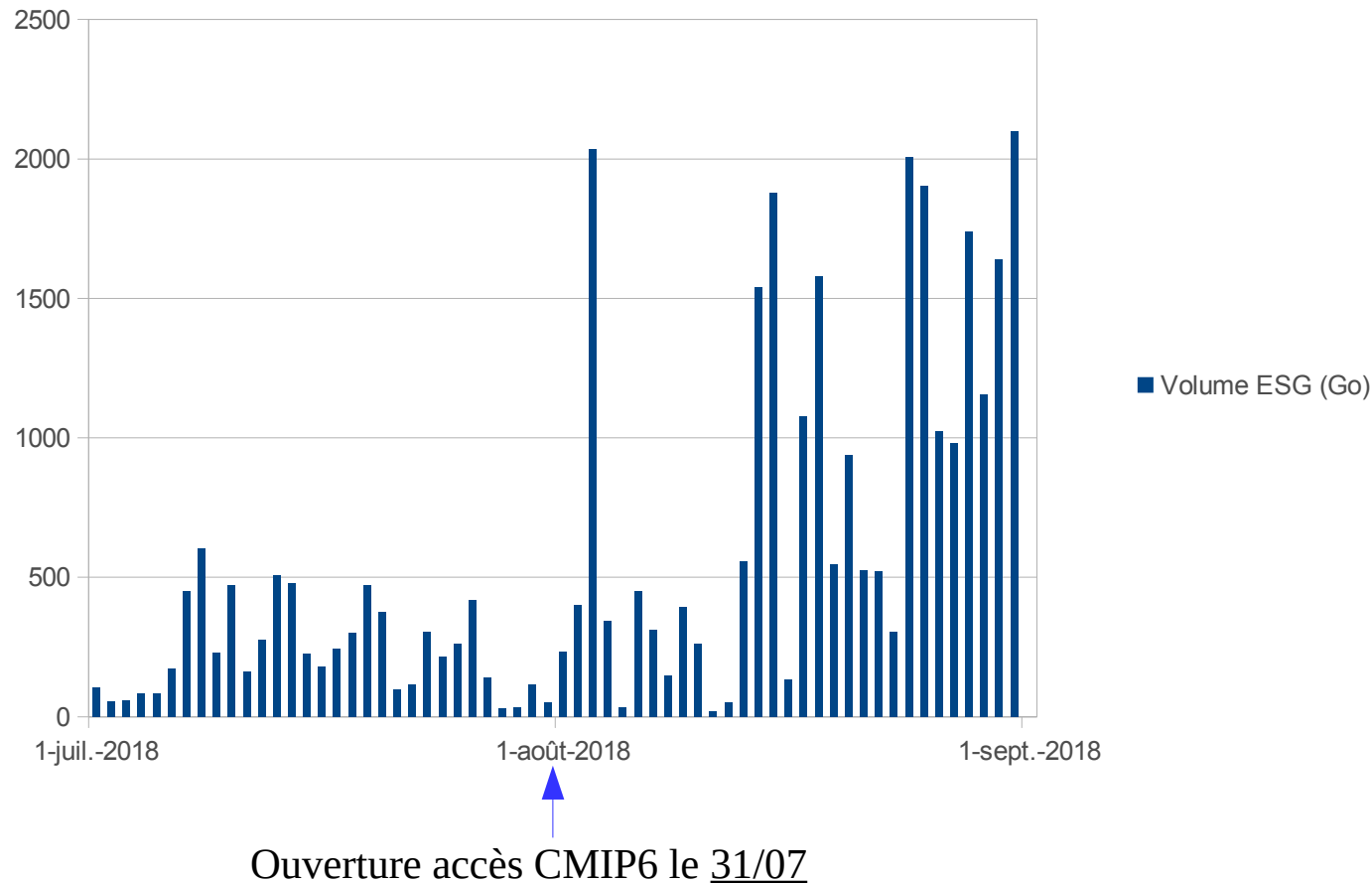
Publication ESGF

- Simulations CMIP6 publiées à ce jour sur l'ESGF seulement par GFDL, IPSL et CNRM (cf liste expériences cr CMIP6-Tech 14/08). Premières publications par NCCS (NASA-GISS) sur fédération de test ESGF
- Evolutions à suivre du composant de publication permettant l'utilisation lors de la publication de tables CMOR de son choix (tables maison incluses)
- Côté CNRM :
 - * publication en cours simulations DECK ESM2-1 'esm-piControl-spinup' et 'piControl-spinup' (post-traitements préalables supplémentaires dûs à bug PrePARE)
 - * publications à suivre : DECK ESM2-1 '1pctCO2', '1pctCO2-bgc' et '1pctCO2-rad'
 - * DataCitation Service (DOI) : utilisation ok de l'API Python proposée par le DKRZ pour obtenir DOI au niveau de l'expérience :
 - pour des raisons de SSI, utilisation de l'API depuis serveur esg-data en DMZ
 - validé sur expérience 1pctCO2, DECK, CM6-1 (autres expériences déjà publiées à suivre) :

<https://cera-www.dkrz.de/WDCC/ui/cersearch/cmip6?input=CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.1pctCO2>

Impact ouverture accès CMIP6 sur le nœud du CNRM

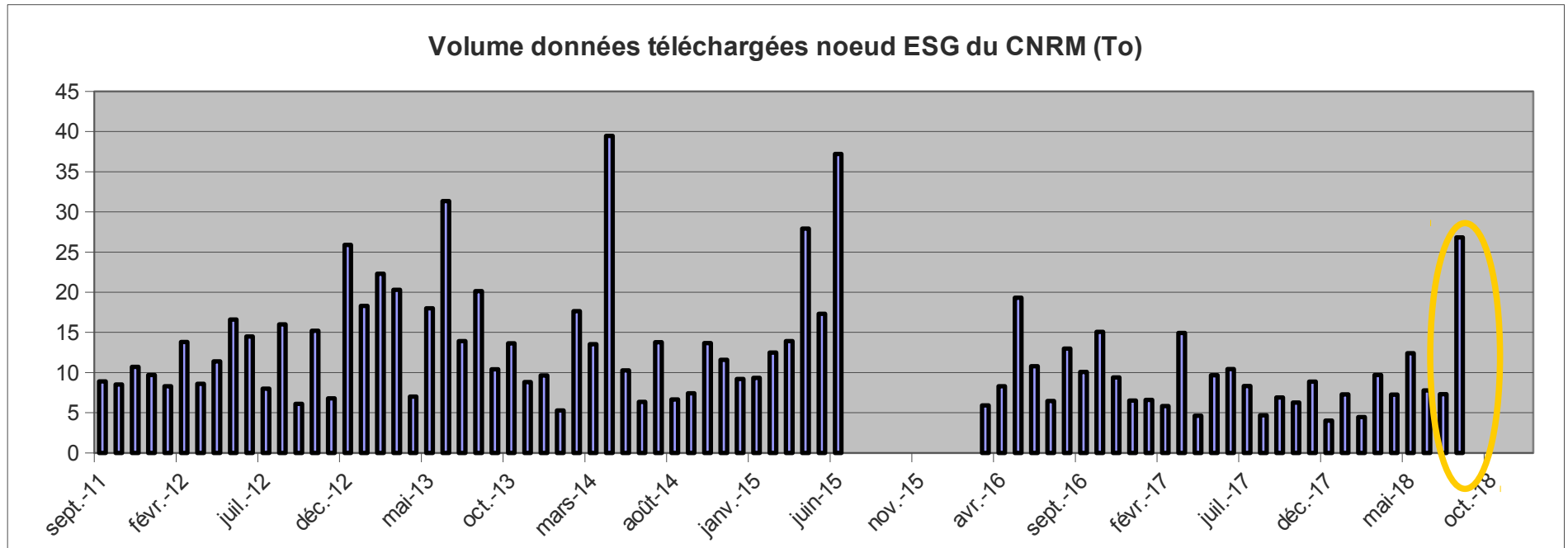
- Effet ouverture accès CMIP6 sur les téléchargements (replica inclus) sur le nœud du CNRM (au 01/09):



Doublement du volume moyen quotidien téléchargé en août (865 Go). Pics >2 To/j

CMIP6-Tech - 13 Septembre 2018

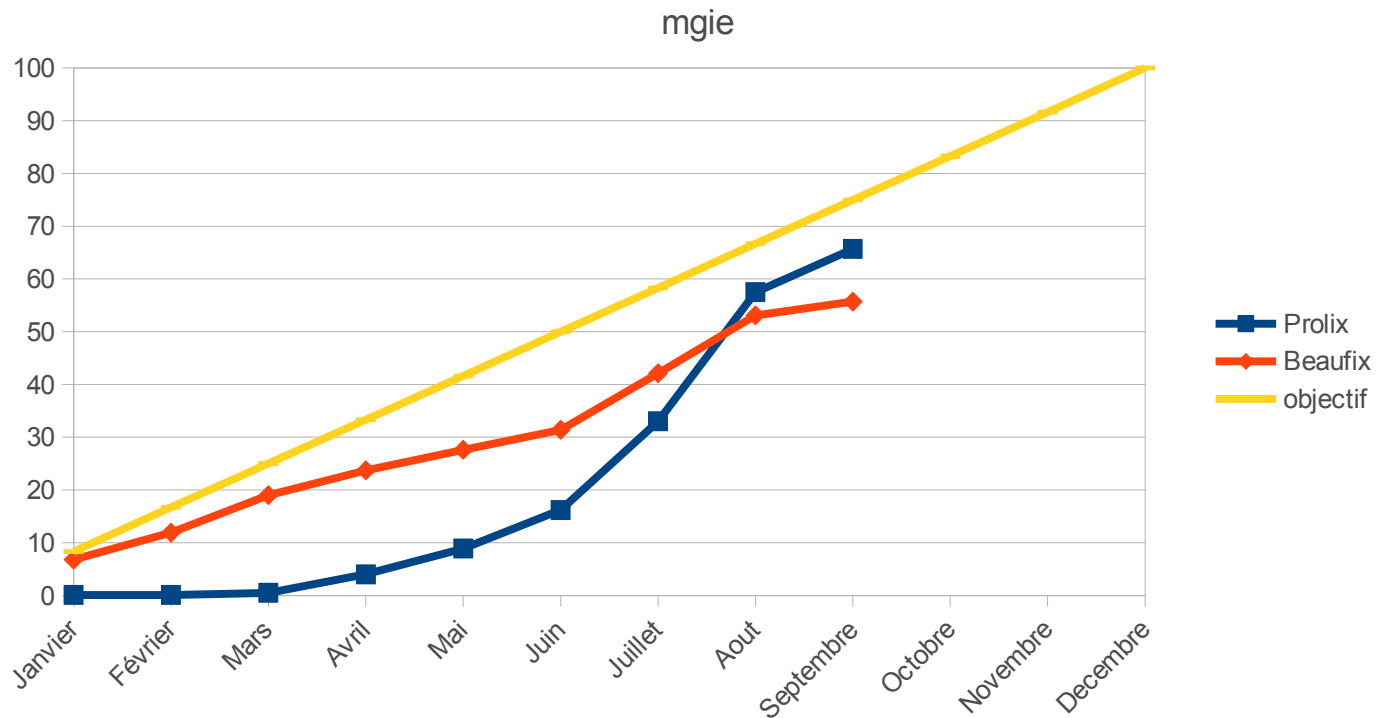
Impact ouverture accès CMIP6 sur le nœud du CNRM



Volume mensuel août 2018 : 5ème volume mensuel le plus important depuis l'ouverture du nœud (26.8 To)

Ressources calcul

- Point consommation heures de calcul mgie (au 13/09) :



Beaufix : 56,5 % quota annuel (reste 38 M d'h), Prolix : 66,3 % (reste 14 M d'h). [Objectif 66,7 % le 01/09]

Eclis et production

- Filière « par job » stabilisé
- Etape finale « assemble_and_QC.sh », prochain Eclis
 - Sauvegarde des fichiers assemblés originaux sur hendrix (sur répertoire original_outputs)
 - Ajout de postpros : ozone, ocnBgchem, branch_time, stdname, history=none
 - Possibilité de rejouer cette étape, depuis des données sur hendrix (et avec reprises sur l'étape ftget)
- Ua7h et consorts :
 - cf mail récapitulatif
 - à intégrer dans assemble_and_QC, par sécurité
- PrePARE sur simulations type piClim_aer ou ssp370-lowNTCF (2 MIPS) :
 - Contournement possible par modif de table (encore à implémenter)
- Evolutions de la DR : impact ? (e.g. ESMVal_baseline)
 - e.g. : expés land-ssp434 et 585 absentes dans DR.21
 - IPSL entame un test de dr2xml sur la DR.27 ; Marie-Pierre a fait aussi un 1^{er} test
- Post-traitements :
 - assemble_and_QC.sh enrichi (version à livrer demain):
 - ▶ fix sur le cas 'no_parent', sur parent_label
 - script de lancement job_postpro.sh
 - ▶ Facile de retraiter les données déjà sur hendrix, générées avec versions antérieures d'Eclis, sans les écraser
 - ▶ A considérer plus tard pour les simus déjà publiées
 - Nota : toutes les erreurs de rectification de meta-données ne sont pas encore « remontées » (ni bloquantes)
- /scratch/work un peu chargé
 - Politique d'effacement des sorties les plus amont à discuter et arrêter !
- Impact des évolutions de la DR : pas encore instruit

```

# Script pour re-traiter des données par assemble_and_QC, pour un peu tous les cas

# Les données d'entrée sont supposées être sur hendrix ou sur scratch,
# à un emplacement conventionnel (déductible du nom d'expérience), et
# seront archivées sur hendrix à l'emplacement conventionnel
# Néanmoins : quand les données sont prises d'hendrix, on ne les écrase pas (on écrit
# les données retraitées dans un répertoire frère : output_postprocessed)

# Un 'user' peut post-traiter les données d'un autre 'user' (mais dans le
# cas de données sur scratch, ca ne marche que s'il demande
# l'assemblage). Ca change la partie 'user' du répertoire de
# destination conventionnel

RUNMAIL=senesi@meteo.fr
EXPID=CNRM-CM6-1_amip-m4K_r1i1p1f2
MIP=CFMIP
user=$(whoami)           # mettre celui qui a fait l'expé, si besoin
user=douvil
from_hendrix=0          # 0/1 - prend-on les données sur hendrix ?
NCRCAT=0                # 0/1 - Veut-on re-faire la phase d'assemblage
LSAVEORIGINALS=1       # 0/1 - Utilisé seulement si from_hendrix=0 et NCRCAT=1. Les
                        # fichiers assemblés sont alors sauvés avant d'être davantage modifiés
PATTERNS=".*"          # utilisé seulement si from_hendrix=1; permet de sélectionner les données;
                        # exemple : "tas_Amon.*"

```


Doc de /cnrm/est/COMMON/CMIP6/get

Get Xios outputs of a CMIP6 run from hendrix and put them in CMIP6 DRS directory structure

You must have ssh working with beaufix/prolix, and ftget working with hendrix (except with option -d)

You must have write permission for shared group cmip6 in order to use the default target location (this is the case for simulanceurs)

You should use only the default value for filter patterns (see below), in order to keep the volume reasonable

Usage : get

-e EXPID experiment to process, with model name prefix and member suffix;
must be set, except with '-d'; must be a CMIP6 one, yet

or

-d DIR don't search hendrix, data is already on directory DIR (-p does not apply, -v is mandatory)
-p PATTERNS filter on data; default corresponds to home data request; can be a series of regexp for matching filenames as e.g. 'tas_.*mon_tos_0mon'; optional; not used with -d
-u USER who did the simulation; defaults to \$(whoami); not used with -d; needed even with -g
-g if data on hendrix should be fetched from 'master account' (mgie002)
-v VERSION data version; by default, it is based on EXPID.conf file modification date; but you cannot process twice the same data with same version number and even with different version numbers, uuid must differ
-r ROOT root location to put data in; default is /cnrm/cmip and is usable only for production run and if you are authroized to write there (OK for simulanceurs)
-f 'ARGS' additional arguments to esgdrs, as e.g. '--set-value activity_id=AerChemMIP' for handling shortcomings in the case of multiple activities (such as piClim-xxx)
-w WORKROOT workdir is experiment specific (this is useful in case transfer is interrupted); When target dir is default dir (see -r), root is on same file system, for efficiency, otherwise, it defaults to ~/tmp; not used with -d
-z for cleaning workdir
-n whatever the case, do Not get files and use only those already in workdir
-x for debug purpose : will do a 'set -x'

For next options, see <http://esgf.github.io/esgf-prepare/drs.htm>

-a ACTION which esgdrs action; default is 'upgrade', other actions are 'tree', 'list', 'todo'
-o OPTION esgdrs default is to move data files; can also use --copy --link --symlink

Examples:

```
./get_CMIP6 -e CNRM-CM6-1_piClim-aer_r11plf2 -u stmartin -f '--set-value activity_id=AerChemMIP' -r ~/tmp/simus  
./get_CMIP6 -d /cnrm/est/USERS/seferian/NO_SAVE/CMIP6/CNRM-ESM2-1_historical_r11plf2/outputs -v 20180701 -o --symlink
```

Production : bon à publier

- Emission du **bon à publier** :

- rappel : renseigner le document GoogleDocs et envoyer un mail contact.cmip@meteo.fr
- préciser la liste des EXPID dont il faut sauver les restarts et listings, pour les cas avec rerun(s)

- Proposition de **mail type** (contribution Roland) :

la simulation CNRM-ESM2-1_esm-piControl-spinup_r1i1p1f2 du MIP DECK est ok pour être publiée.

les outputs et restarts sont localisés sous: `hendrix:/home/seferian/CMIP6/DECK/CNRM-ESM2-1_esm-piControl-spinup_r1i1p1f2/iox/output`

Les listing de prepare et nctime sont sous: `seferian@sxest:/cnrm/est/USERS/seferian/NO_SAVE/CMIP6` (normalement : sous \$RELDIR)

Cette simulation a subit le check-up suivant pour les erreurs connues: (*pour certaines, à partir de la 4^e, on peut mettre : « OK – Eclis-6.38 »*)

- 1hrClimMon : sans objet pour cette simu (*pour les autres, pour l’instant, indiquer ‘reste à traiter’*)
- correction de sftlf : corrigé (si version CMIP6/V2, indiquer ‘sans objet pour V2’)
- pas de temps des CFsubhr : non concerné

[Les compte-rendus ci-dessous peuvent être remplacés par : « post-traité avec Eclis 6.38 »]

- nommage des variables ua7h et al. (6hrPlev : hus4 et wap4 // 6hrPlevPt : hus7h, ta7h, ua7h, va7h, wbtemp7h, zg7h // Omon : ficeberg2d) : conforme
- unités de l'ozone : pas nécessaire avec CNRM-ESM2-1
- orthographe de ocnBgchem : corrigé
- correction des standard_names : non
- correction des parent_variant_label (pour les expés ayant un parent) : sans objet
- ajout des attributs du parent pour les expés sans parent : oui/non

Les fichiers originaux sont conservés sous: `hendrix:/home/seferian/CMIP6/DECK/CNRM-ESM2-1_esm-piControl-spinup_r1i1p1f2/iox/output/original`

Liste des variables ou Tables à exclure de la publication : AUCUNE

Analyses

- CliMAF :
 - Accès à des données via ftp OK
 - Tjs à voir : CliMAF pourrait-il gérer plusieurs caches (en lecture), dont un cache partagé qui serait écrit par un/des users privilégiés ?
- C-ESM-EP
 - Le trunk convient pour tourner au CNRM (merge fait)
 - Yapuka : Documenté sur site C-ESM-EP : <https://github.com/jservonnat/C-ESM-EP/wiki/Specifics-for-using-C-ESM-EP-at-CNRM>
 - 'merge' des versions CliMAF : reste à faire