

Réunion Tech

- production CMIP6 – partie amont
 - publications CMIP6
 - ressources et site Web
 - Eclis 6.40
 - transferts
 - ressources ESPRI
-

08 Mars 2019

Orfévrerie de la production CMIP6

- Cf page suivante pour la discussion et les décisions sur les variables « à problème »
- TasLut, hussLut, mrsosLut sont erronées à cause de dr2xml. Récupérable par post-traitement
- TasLand ne comporte par l'influence de la neige. Incurable. Question : du coup, est-ce que tas peut être correcte
- Un certain nombre de variables seaice n'ont pas de pondération par sic dans la moyenne temporelle ; dont siu, siv. Incurable. Mals les données son utilisables ; et on continue comme ça.
- Similairement pour qlq variables « isccp » (Romain compte re-traiter) et land (tsn, agesno : Bertrand OK)
- Le commentaire de description de la grille de Trip comporte un caractère spécial → pb pour la réplication seulement. Un palliatif est en cours de déploiement sur les data nodes → pas d'action
- Les 'bounds' des coefficients 'a' et 'bp' sont transposées. Décision : indication d'une known issue et fourniture d'un fichier correct de ces champs fixes
- F=2 : demande faite auprès du WIP pour changer le commentaire prévu (qui n'est pas encore apparu)
- Assemblage des fichiers HR par un filière basée sur la taille
 - On donne une taille comme valeur de SPLIT_FREQ
 - Déboguée par Aurore
 - Problème sur les temps d'exécution : il faudrait que assemble_and_QC.sh puisse être multi-noeuds, ou qu'un découpage intervienne en amont
- Calculs de climatologies et de moyennes décennales : ne seront pas traités

Discussion et décisions sur les variables 'à problème'

- Chaque 'responsable de variable' doit statuer pour une variable à problème, sur:
 - Variable utilisable ou pas (ex : 'siu' est utilisable même si la pondération n'est pas la bonne). Si non, il faudra dépublier
 - Si pas utilisable, intérêt à recalculer, et modalités
- Dans le dernier cas, chaque responsable de MIP statue sur celles des expés qui donnent lieu à recalcul
- Ces éléments sont examinés en réunion Tech. On tâche d'éviter les situations 'patchwork' où des données seraient justes dans certaines simulations et 'moins justes' dans d'autres
- Nous ne publierons pas d'errata concernant des simulations nombreuses (sauf si le système d'errata le facilitait) ; mais nous renseignerons la page 'known issues' (encore à construire) de notre site CMIP6
- Plus globalement, le mode de traçage de tout problème remonté doit être instruit en listant les différentes options de traitement possibles : simple traçage en interne et/ou externe puis traitement par simple alimentation d'une page de type 'Known issues' ou par alimentation de l'Errata Service ES-DOC <https://errata.es-doc.org/static/index.html> avec actions de correction ou non (actions possibles : dépublication ou non, retraitement des datasets concernés ou non, republication éventuelle).

dr2xml : Point d'information

- Branche gr_refactoring (disponible sur le dépôt) : modification de la structuration de dr2xml
- Correction envisagées en plus :
 - Problèmes avec les variables tasLut, hussLut et mrsosLut.
Liste des variables avec une dimension singleton hauteur et un autre axe ?
Cas où il faut enlever des axes ?
 - Prise en compte des included_variables dans le dictionnaire simulation_settings :
à la place ou en plus de celles du dictionnaire laboratory_settings ?
- Ajout de variables sans modification à la main des fichiers xml : correction pour l'utilisation des variables « perso » et première proposition pour les variables « dev » :
 - Dans la DR maison :
perso; titi ; ocean; month; CNRM_month; time-mean; XY-na; ANY; ANY
dev; hmv1; ocean; day; CNRM_day; time-mean; XY-na; ANY; ANY; my_units; my_long_name;
my_std_name; grid_T; grid_T_2D
 - Version en développement de dr2xml associée (beaufix) :
/home/mf/dp/marp/rigoudyg/dev_dr2xml/dr2xml_simple_variables
 - **Idée évoquée : qu'il existe un axe de pressions de nom fixe, mais de niveaux configurable dans les settings**
- Nouvelle version « numérotée » vers le 20 mars avec toutes les corrections/évolutions disponibles ?
- **Le rôle premier du Cerfacs dans la gestion de la doc de dr2Xml est confirmé**

Publication ESGF : informations générales

- Publication CMIP6 réalisée à ce jour sur 9 nœuds. Détails sur l'ESGF et sur la page de monitoring :
https://pcmdi.llnl.gov/CMIP6/ArchiveStatistics/esgf_data_holdings/
- Réplication CMIP6 : 100 To d'ores et déjà répliqués à l'IPSL, 35 To à NCI. Données CNRM-CERFACS en partie répliquées à l'IPSL et au LLNL
- Toujours aucune nouvelle de la communication envisagée pour l'ouverture de l'accès aux simulations CMIP6 au niveau international...

Publication ESGF au CNRM

■ **Bilan au 07/03 :**

- 200 bons à publier émis dont 36 simulations DECK (19 CM et 17 ESM)

- 53 simulations publiées (~150 To) :

* fin publication DECK (CM et ESM) le 05/03

* fin publication simulations Crescendo avec jalon mars [AerChemMIP et RFMIP] (reste piClim-lu en cours de publication)

* début publication ScenarioMIP CM6-1 : r1 ssp126 CM6-1 publié le 07/03, r1 ssp245 en cours

* rappel fichier suivi publication (utilisé pour tableau de bord site web CMIP6) : fichier SuiviPubliGlobal.ods sous /cnrm/est/USERS/franchisteguy/CMIP6/CNRM/Publication/SuiviPubli

■ **A venir** (par ordre de priorité décroissant, **validé en réunion**) :

* membres r1 ScenarioMIP CM

* DAMIP (hist-aer, hist-GHG, hist-nat)

* membres r1 ScenarioMIP ESM

* omip CM (après déclaration CMIP6 CV) et ESM

* autres membres ScenarioMIP CM et ESM

■ ES-DOC : fourniture d'éléments descriptifs des modèles à faire avant le 31/03 (puis demande autres éléments à suivre ensuite). Initialisation fichiers à partir des éléments CMIP5 (A/R avec équipe ES-DOC) à valider et compléter. Visibles sur GitHub <https://github.com/ES-DOC-INSTITUTIONAL/cnrm-cerfacs> . Initialisation très décevante pour le moment !! **Nous convenons de ne pas essayer les plâtres**

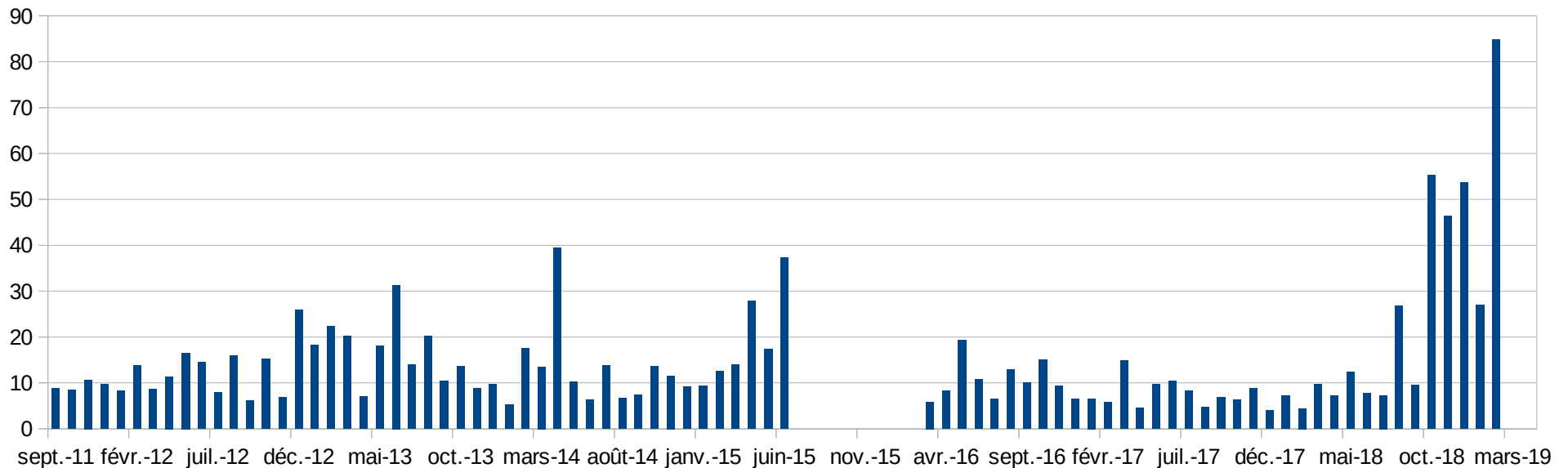
■ Echanges en cours avec IPSL sur les conditions nécessaires pour l'activation de gridFTP et Globus sur le nœud du CNRM

■ Publication du Cerfacs sur le nœud du CNRM : actions à mener listées avec Marie-Pierre et CTI. Premiers tests de génération de mapfile réalisés au Cerfacs, tests de publication à suivre

Noeud ESGF au CNRM

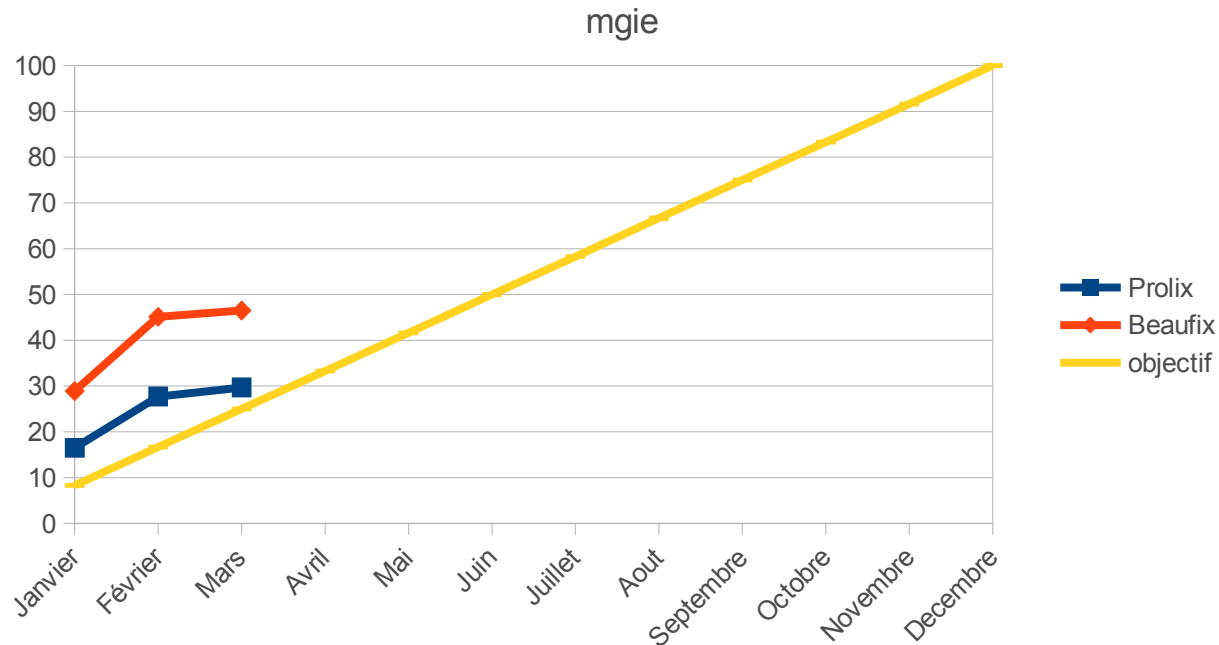
- Records quotidien et mensuel volumes téléchargements battus en février sur le noeud :
 - quotidien : 10.9 To le 24/02
 - mensuel : 84.8 To (ancien record 55.3 To en octobre 2018)

Volume données téléchargées noeud ESG du CNRM (To)



Ressources calcul et stockage, site web

- Ouverture accès externe site web CMIP6 <http://www.umr-cnrm.fr/cmip6/> le 20/02 (référencement Google ok)
- Consommation calcul mgie (au 07/03) : Beaufix (46.5 %), Prolix (29.7 %), 4 Mh sur queue Libre Beaufix depuis début 2019
- Il faut appeler les équipes à ne pas sur-utiliser l'account mgie pour des simus qui n'étaient pas initialement prévues, et ce même en queue libre. A noter que les équipes sous-consomment sur leur account propre.



Eclis V6.40

- DATRES et DATSFXRES obligatoires, pour spécifier les binaires de changement de la date dans les restarts initiaux pour Arpege et Surfex.
- UPOSTPRO désigne un script de post-traitement en mode LIOXSAFE=1
 - \$ECLIS/toolbox/postpros_CMIP6.sh;
 - \$ECLIS/toolbox/fix_o3_CNRM-CM6-1.sh qui permet de pallier le problème de facteur correctif pour l'ozone 3D d'Arpege 6.4.1 en version GCM.
 - ... Cordex
- HDL pour l'UUID dans les fichiers assemblés
 - "hdl:\${HDL}/\${(uuidgen)"; la valeur par défaut est celle qui va bien pour CMIP6
- LSAVE_BEFORE_POSTPROS=0 si on a confiance dans UPOSTPRO
- nettoyage sur d'anciennes fonctionnalités (report de TS Surfex dans les restarts Arpege, vieux formats de Trip et de Surfex, CREDAS)
- les messages sur plantage conseillent désormais d'utiliser brelan pour les relances ; xrelan est conservé mais pas vraiment recommandé
- nrcat_dir_by_size.sh débogué
- Vérifier qu'il y a aussi l'option 'carry_on' de job_postpro.sh (proposée par M-Pierre)

Transferts

- /scratch/CMIP6 : expliqué, transféré
- Dr2xml :
 - Dépôt transféré à Gaëlle
 - Restructuration du code : validé ; mise à disposition
 - Traitement facilité des variables nouvelles
 - Explications faites à l'IPSL sur « comment respecter la norme CF pour les coordonnées hybrides »
- Eclis
 - Une V 6.40 en cours de merge (sans les runs annuels)
 - Répertoire de références chez Gaëlle jusqu'à V6.39
 - Ne plus mettre de nouvelle version sur /scratch/CMIP6 [après la V6.40](#)
- CliMAF :
 - Présentation de l'architecture à Gaëlle et Jérôme sur 8 séances
 - Nombreux points d'amélioration identifiés
 - Axes de développements discutés
 - Transfert de propriété du dépôt GitHub à suivre.

Ressources ESPRI (mail David Salas)

- j'ai une question portant sur vos besoins sur la plate-forme ESPRI (machines Climserv + Ciclad).
- Il m'a été rapporté que certains commencent à se sentir un peu à l'étroit dans l'espace alloué (de l'ordre de 100 Go/personne), c'est peut-être le cas aussi en termes de puissance de calcul (mais plus difficile d'y trouver une solution). Cela m'incite à faire le nécessaire auprès de l'IPSL (en termes de financement) pour arranger cette situation.
- Il me faudrait cependant connaître un peu mieux les besoins. La réponse est peut-être simple "doublement du stockage pour tous" (vous pouvez en discuter un peu en réunion), ou un peu plus élaborée : je peux revenir vers tous les utilisateurs actuels, mais aussi identifier de nouveaux utilisateurs (nouveaux CDD et nouveaux permanents) au moyen d'un nouveau sondage.
- Pour le moment, j'ai demandé à l'IPSL de me fournir une estimation des coûts de l'utilisation d'ESPRI (dont espace, CPU...) car c'est, avec l'expression de vos besoins un des éléments me permettant de bien dimensionner le financement que je dois apporter.
- **Sujet pas traité faute de temps (tour de table en réunion et pas de problème particulier à remonter)**
- **Prochaine réunion le mardi 2 avril à 14h30 salle 151**