

Réunion Tech

- CliMAF
 - C-ESM-EP
 - Eclis
 - dr2xml
 - Production CMIP6
 - Publications CMIP6
 - Ressources et site Web
-

24 septembre 2019

- Documentation et formation :
 - Les baux CliMAF reprendront en septembre, mais les listes de diffusion fonctionnent toujours en cas de problème.
 - La documentation et les exemples nécessitent une mise à jour (ex : utilisation de crm avec des pattern)

- Retour suite aux discussions pendant l'AG GMGEC :
 - Formalisation des étapes pour la soumission des contributions à CliMAF :
 - ▶ Côté Toulouse, les contributions sont présentées en réunion GMGEC-Tech (mail préalable pour donner les diapo correspondantes)
 - ▶ Si elles sont validées, transmission des contributions avec la documentation associée, vérifications que les modifications n'ont pas d'effet de bord puis intégration
 - Problème récurrent du pré-traitement des données nécessaire avant utilisation dans CliMAF pour tous les fichiers un peu exotiques.

CliMAF : Dernière version disponible (V1.2.12)

- Contenu :
 - Modifications pour le Cerfacs et l'IPSL
 - Modifications relatives à la C-ESM-EP
 - Corrections de bogues (dont lecture des champs fixes)
 - Ajout de fonctions permettant de ne pas refaire de sélection avec cdo (devrait résoudre en partie les problèmes de lenteurs) : [faire la liste de ces opérateurs *_fast](#)
 - PEP8
- Compatible avec la C-ESM-EP (v1 et v2)
- Disponible depuis le 11 septembre.

- Bogues connus :
 - Problème dans certains cas avec sellonlat : lié à cdo
 - Lenteurs dans les cas où les fichiers sont gros : utilisation des fonctions *_fast à généraliser
 - Problème dès que les calendriers utilisés ne sont pas standards : lié à la fonction check des dataset et aux comparaison de dates
 - [Problème de lecture d'un champ fixe dans un projet \(voir si cela fonctionne avec fds plutôt qu'avec la définition d'un projet\)](#)
 - [Débogage nécessaire de la fonction check des dataset](#)

CliMAF : Développements

- V2.0 :
 - Compatibilité Python 3
 - Renommage de la fonction slice
 - Mise à jour de la doc et des exemples
 - Rédaction de tests unitaires
 - Vérifier le fonctionnement de la création d'objets sans utilisation du cache
- Et pour plus tard :
 - Arrêter d'utiliser glob.glob
 - Implémenter une méthode de suppression en fonction de la durée de création
 - Simplifier et déboguer les projets *_extent (récupération d'un historique sur une période, d'un scénario pour la suite)
 - Homogénéiser la sélection des données (ds, cdataset, cens, eds, ...)
 - Remplacer les scripts NCL par des scripts python (voir pyngle) :
 - Exemple de script déjà disponible sur lustre (tracé de trois cartes superposées : couleurs, contours, hachures) :
`/cnrm/est/USERS/rigoudyg/dev/scripts_plot/plotmap.py`
 - Interrogation du service d'errata
 - Récupération de données sur des sites distants
 - [Regarder s'il est utile/envisageable d'appeler cdo depuis le module python dédié](#)

Atlas et C-ESM-EP

- C-ESM-EP :
 - Version V2.0 de la C-ESM-EP presque finalisée par Jérôme (devrait être officiellement disponible rapidement)
 - Atlas réalisés par Laure Coquart sur les simulations CMIP6 disponibles sous : <http://cerfacs.fr/giec6/>
 - Version disponible sur le lustre du CNRM sous /cnrm/est/USERS/rigoudyg/C-ESM-EP.
Utilisable après copie chez soi.
Les jobs sont lancés sur la partition P8HOST sur aneto (valable pour tous les serveurs).
 - Vérifier qu'il est possible de mettre l'option adéquate dans les plot pour le régional :
 - ▶ Ne donne pas des résultats satisfaisants à première vue (données Aladin projetée sur la grille des observations et non l'inverse)
 - ▶ `file:///cnrm/est/USERS/rigoudyg/NO_SAVE//C-ESM-EP/standard_comparison_aladin_rigoudyg/Atmosphere_Surface/atlas_Atmosphere_Surface_standard_comparison_aladin.html`
- Atlas / Générateur automatique de diagnostics :
 - Développements toujours en cours
 - [Réunion vendredi 27/09 sur le sujet](#)

Eclis : Versions

- V6.43 (depuis le 9 août) :
 - Augmentation du realization_index
 - Correction du plugin atm_lbc_nudging pour pouvoir commencer des run Aladin en milieu de mois
 - Corrections de bogues
 - Correction de cm_files
 - Changement de la version de NCTIME par défaut : 4.6.4 (correction du bogues liées aux dates supérieures à l'an 4000)
 - Ajout de variables disponibles via le fichier param pour gérer les caractéristiques du job assemble_and_QC : ELAPSAQC, ELAPSFRONTAQC, MEMAQC, MEMFRONTAQC, QOSAQC, QUEUEAQC
- V7.0
 - Poursuite des développements
 - Développements des branches V6.* et V7.0_beta* remontés régulièrement
 - Début du travail sur le nudging avec récupération des différentes versions, préparation pour les run aladin
 - Vérifier que le nombre de digit est paramétrable : [c'est ok](#)
 - Vérifier la durée maximale faites dans un run unitaire (a priori au moins 2 ans) : [pas de problème dans Eclis, mais plutôt dans UPDCLIM](#)

Eclis : Prochaine version (V6.44)

- Développements finalisés :
 - Ajout d'un flag pour indiquer s'il s'agit d'une simulation de production et pour quel projet
- Développements en cours :
 - Finalisation de job_postpros_new et assemble_and_QC_new (utilisation de mtool)
 - Test changement de version de mtool
 - Mettre un warning dans relan si jamais il y a déjà eu un post-traitement pour éviter d'avoir des données plusieurs fois sur Hendrix
- Récupération des développements en cours sur la V7.0 concernant le post-traitement :
 - Réécriture d'une partie des scripts utilisés par le post-traitement en python pour plus de flexibilité
 - Ajout de la possibilité de faire une partie des post-traitement avant la concaténation des fichiers (exigence HR utilisable pour d'autres simulations)
 - Ajout de possibilités pour la concaténation par taille

Eclis : Et pour la suite...

- Développements prévus :
 - Suppression de LDR_EXPID
 - Interfaçage de xiofilechecker (vérification de l'adéquation des fichiers xml et des fichiers produits)
 - Interfaçage d'un débogueur (voir la documentation d'Aurore sur le sujet)

- Développements à discuter à l'occasion :
 - Vérifier quelles variables devraient être modifiées par Eclis dans les settings de dr2xml (en lien avec les développements dr2xml)
 - [Intégration d'un équivalent de verifrelan dans Eclis ? Voir si les fonctionnalités de slurm le permette comme on ne peut pas faire de cron sur le calculateur. A discuter en fonction.](#)
 - [Les post-traitement peuvent se faire au fil de l'eau, voir s'il faut le généraliser au vu de la charge de lustre sur les calculateurs.](#)

- Documentation :
 - Mettre à jour la documentation sur la façon de faire des runs avec Eclis et dr2xml
 - Communiquer sur l'utilisation de LDR_EXPID et expid_in_filename (problèmes avec différents outils lors de l'utilisation par la suite : get, CliMAF, C-ESM-EP...)
 - Mettre à jour les param de référence sous */scratch/CMIP6/V2*

dr2xml : Versions

- V1_CMIP6_dev :
 - Version parallèle qui embarque toutes les nouveautés des branches récentes sans les modifications pouvant impacter le contenu ou les métadonnées des fichiers
- V2.0 (3 juin 2019) :
 - Compatibilité python 3
 - Axe de pression configurable
 - Fichier contenant les pattern des variables dev et perso
 - Possibilité de réduire les méta-données
- V2.1 :
 - Disponible dès qu'elle sera deboguée
 - Contenu : cf diapo suivante

dr2xml : Version en développement (V2.1)

- Développements finalisés :
 - Adaptation pour Cordex : V2.0_xml_cordex
 - Il reste à voir l'ajout des objets de type char (de la même façon que pour NEMO)
 - Gestion des caractères spéciaux :
 - Il faut encoder les fichiers contenant des caractères spéciaux en utf-8 (et même tous les fichiers texte :-)) :
`iconv -f encodage_initial -t utf-8 fichier_init -o fichier_out`
 - Gestion interne de la lecture et de l'écriture des fichiers xml avec conservation de l'ordre.
 - Début de réalisation d'une doc sphinx
 - Correction de l'absence de l'attribut `branch_method` pour les simulations n'ayant pas de `parent_id`
- Reste à faire :
 - Corriger un problème lors des lancements avec Python 3 sur certaines simulations
- Mise en place d'une base de tests permettant de vérifier la reproductibilité entre deux versions du code et de Python

dr2xml : Et pour la suite...

- Généralisation de dr2xml à d'autres projets (en fonction des demandes) :
 - Actuellement, présence de tests pour vérifier si on se trouve dans le cas Cordex ou dans un CMIP6).
 - Utilisation d'un fichier de settings propre au projet
 - Nécessite de savoir quelles variables des settings dr2xml sont propres à CMIP6
 - Généraliser la structure pour qu'elle soit plus adaptée aux différents projets
- Utilisation d'un fichier settings propre aux variables dev ?
 - Ajout des axes personnalisés
 - Gestion des grilles natives pour les variables dev
- Poursuite de la rédaction de la doc
- Simplification de l'installation (pip, ...)

Production CMIP6

- Liste des champs fixes manquants à finaliser et à générer si besoin.
- Variable vt100 :
 - Version 1.00.31 de la data request disponible avec des modifications sur la variable vt100 mais a priori pas finalisées
- Définition des bassins pour le HR ?
Point mis de côté, personnes n'a les bassins.
- Simulations HR en cours, post-traitement :
 - Le DECK est terminé
 - La moitié de la simulation amip est terminée
 - Les post-traitements sont en cours (quelques problèmes liés au HR détectés)
- 4 simulations LS3MIP mal définies dans la DR (pas de nom différent pour les simulations quand les scénarios sont différents). Décision unilatérale de leur côté pour f1 → ssp126 / f2 → ssp585.
Problème utilisation de f2 (sauf exceptions ou f>2).
Préparation d'un mail pour indiquer que les indices de forçages qui seront utilisés ici seront f2 pour ssp585 et f21 pour ssp 126.

Publication ESGF : informations générales

- Publication CMIP6 réalisée à ce jour sur 23 nœuds. Détails sur l'ESGF et sur la page de monitoring :
https://pcmdi.llnl.gov/CMIP6/ArchiveStatistics/esgf_data_holdings/
- Nouvelles versions de la pile logicielle ESGF (v4.0.3 et v4.04) :
 - * montée de niveau réalisée à l'IPSL les 10 et 11/09 (v4.04)
 - * installation 4.0.3 sur le nœud de test du CNRM effectuée le 16/07, tests de publication et de robustesse ok
 - * à suivre au CNRM : montée de niveau 4.0.3 sur le nœud de production
- Départs IPSL : S. Gardoll (point de contact technique pile ESGF, 15/10) d'ores et déjà remplacé par un CDD
- Pas de réunion du CDNOT depuis le 26/06 (chairman en partance....)

Publication ESGF au CNRM

■ Bilan publication CNRM au 23/09 :

- 226 bons à publier émis (+ 3 extensions 2100-2300 membres r1 ScenarioMIP ESM2-1)

- 220 simulations publiées (~295 To) soit ~75 % du nombre total de simulations à publier sur l'exercice. 13 MIPs dorénavant couverts côté CNRM-CERFACS

- Depuis la dernière réunion Tech (06/08) :

* fin publication simulations C4MIP, GMMIP

* complément 'piClim-control' ESM2-1

* début publication LS3MIP et poursuite publication CM6-1 CFMIP (piSST*, a4SST*, amip-*)

■ Echecs de publication sur plusieurs simulations dus à l'absence de l'attribut global 'branch_method' → post-traitements nécessaires

■ A venir côté CNRM (par ordre de priorité décroissant, à valider) :

* dépublication amip-piForcing et aqua-p4K (voir point 'known issues')

* attente feu vert David pour publication tables Slday et Slmon (68 simulations CNRM en attente)

* simulations CM6-1 CFMIP (3 amip-* restants)

* extension '2300' membres r1 ScenarioMIP ESM2-1

* esm-hist r1i1p1f3 (modèle en émission)

* complément DECK piControl LR (500 premières années seulement publiées) et simulations DECK HR après émission bons à publier correspondants

* puis ordre chronologique d'arrivée des futurs bons à publier

Activity	
<input type="checkbox"/>	AerChemMIP (14332)
<input type="checkbox"/>	C4MIP (2822)
<input type="checkbox"/>	CFMIP (3736)
<input type="checkbox"/>	CMIP (16242)
<input type="checkbox"/>	DAMIP (5161)
<input type="checkbox"/>	DCPP (20919)
<input type="checkbox"/>	GMMIP (1013)
<input type="checkbox"/>	GeoMIP (2038)
<input type="checkbox"/>	LS3MIP (382)
<input type="checkbox"/>	LUMIP (1202)
<input type="checkbox"/>	OMIP (143)
<input type="checkbox"/>	RFMIP (1858)
<input type="checkbox"/>	ScenarioMIP (20938)

Nb dataset CNRM (23/09)

Publication ESGF au CNRM

■ Côté CERFACS :

* publié : DCPD amv-pos (40 mbrs), amv-neg (40 mbrs), ipv-pos (25 mbrs), ipv-neg (35 mbrs)

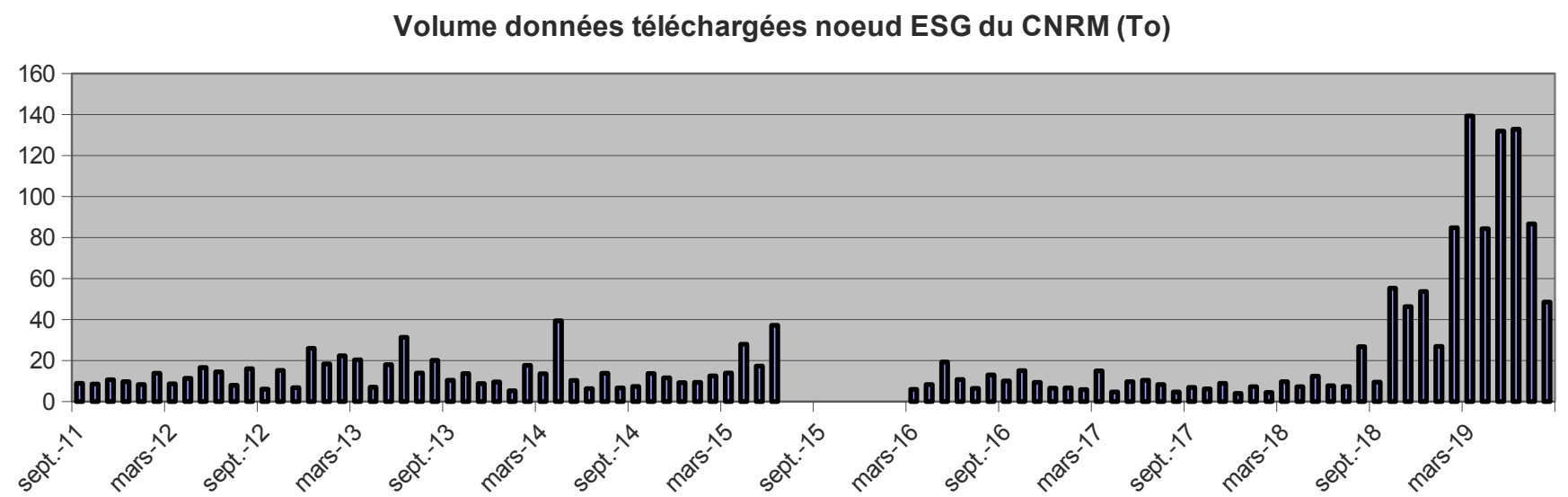
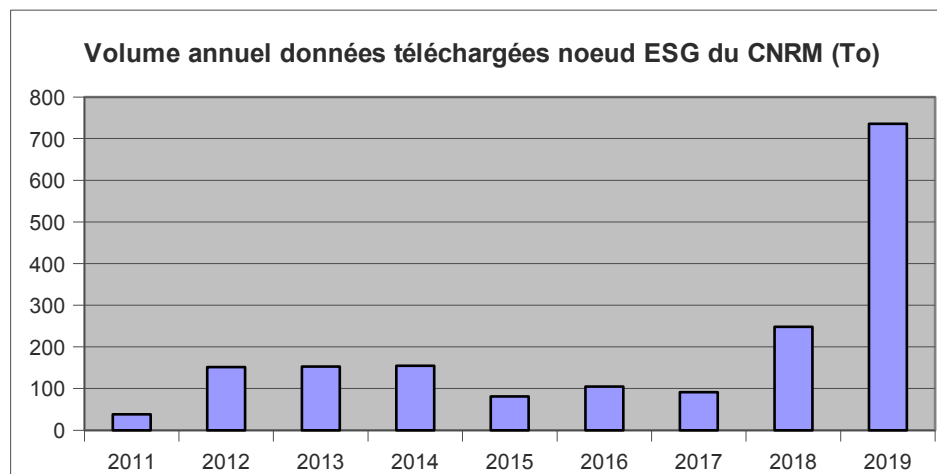
* à venir :

- membres additionnels des amip (présent et futur, 9 mbres supl. chacun), des historiques (3 mbrs supl.), des scenarios (3 mbrs supl.) de HighResMIP

- PAMIP (plutôt en octobre/novembre)

Trafic noeud ESGF au CNRM

- Record volume quotidien maximal téléchargé battu sur le noeud : 11.3 To le 28/07 (ancien record 10.9 To au mois de février 2019)
- Volume total provisoire 2019 téléchargé (~750 To) équivalent au volume téléchargé sur le noeud sur la période 2011-2016
- Baisse récente volumes téléchargements due à un problème sur fonctionnalité wget sur notre noeud (résolu)



Errata Service, known issues

- variables **seai** mal pondérées (cf mail Stéphane 23/04 : siv, siu, sisnconc, sisnmass, sisnhc, sistrxdtop, sipr, sistrydtop, sistryubot, sistrxubot, sidragbot, siflwdrain, siflwbob, siflswutop) : attente verdict final (David) pour décision de dépublication (738 datasets déjà publiés) et publication jeux à suivre. Publication tables Slday et Slmon stoppée depuis le 16/04 dans l'attente de cette décision
- variable **cfadDbze94** buguée (cf mail Romain, déclaré sur errata-service et sur le site web CMIP6 CNRM le 22/08)
- attribut **parent_variant_label** erroné pour certaines simulations ('variant_label' au lieu de 'parent_variant_label'). A suivre : liste des simulations concernées, texte pour site web (rubrique 'known issues')
- simulations CM6-1 et ESM2-1/historical (a minima) : problème sur l'axe des temps : contenu variable **time** pour **00h** : "YYYY-MM-DD" au lieu de "YYYY-MM-DD 00" pour tables 6hrLev et 6hrPlevPt. A suivre : à discuter. [Contenu logique \(idem CMIP5\), les fichiers pointés sont en fait corrompus et seront recalculés \(dépublication/republication à faire\)](#)
- 2 simulations à dépublier (+déclaration Errata Service + texte site web) :
 - simulation **amip-piForcing** CM6-1 (CFMIP) publiée mais période incomplète (1979-2014 au lieu de 1870-2014). Rerun à suivre par Romain
 - simulation **aqua-p4K** CM6-1 (CFMIP) publiée mais réalisée avec config aqua-p4K-lwoff. Rerun à suivre par Romain
- Rappel : tableau (usage interne) recensement erreurs sous [/cnrm/est/USERS/franchisteguy/CMIP6/CNRM/Production/Issues \(InfoIssuesInterne.odt\)](#)

Errata Service, known issues

- Problème avec les simulations r1i1p1f2 des Scénarios ssp585 et ssp370 de l'ESM qui ne sont plus reproductibles :
 - Pour ssp 585, problème de fichier LandUse corrigé depuis
 - Pour ssp 370, problème de fichiers corrompus
- Les simulations seront dépubliées (fait le 27/09) et refaites complètement.
- Pour les simulations qui sont parties de ces scénarios :
 - Pour les extensions, simulations à vérifier (normalement juste effet papillon)
 - Pour les autres, confirmer que les simulations sont à refaire

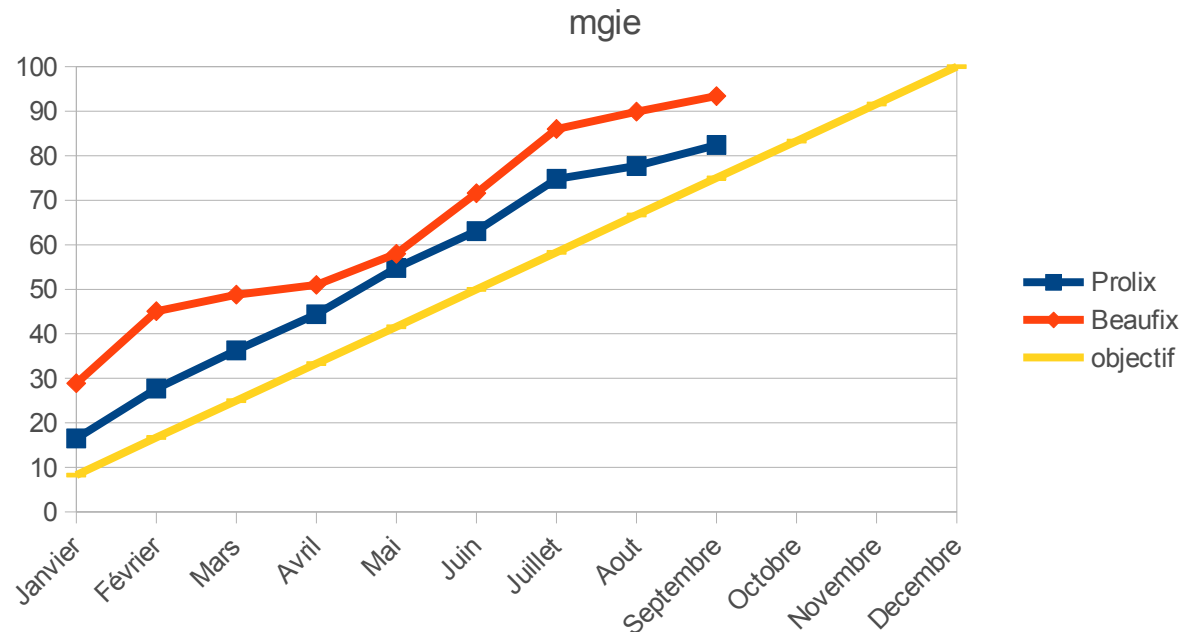
- ESDOC alimenté à ce jour par 10 centres, cf <https://explore.es-doc.org/cmip6/models>
- Coté CNRM, pour ESM2-1 (cf mail de relance 17/09) :

	landice	aerosol	atmoschem	land	coupling	atmos	ocean	ocnbgchem	seaice	toplevel	citation
Fichier xls	-	Ok	Ok	Ok	Ok	Ok	Ok				
EsDoc	-	Ok	Ok	Ok	A publier	A publier					

- Toujours en attente de réponse de l'équipe esdoc sur certaines issues (cf <https://github.com/ES-DOC-INSTITUTIONAL/cnrm-cerfacs/issues>)
- A suivre :
 - 1) récolte puis vérification des fichiers restants pour esm2-1 (ocean, ocnbgchem, seaice, toplevel et citation), dépôt sur GitHub ESDOC et envoi à l'équipe ESDOC de vos remarques éventuelles sur les fichiers à renseigner
 - 2) extension ensuite à cm6-1 puis cm6-1-hr

Ressources calcul et stockage, site web

- Rappel renouvellement calculateurs : arrêt Beaufix/Prolix prévu le 20/09/2020
- Consommation calcul mgie (au 23/09) : il ne reste que 2.1 Mh sur Beaufix et 2.4 Mh sur Prolix sur le quota 2019 (NB : environ 15 Mh déjà consommées en queue Libre sur Beaufix sur mgie en 2019)
- Nette sur-consommation et quota théorique insuffisant pour finir ce qui est prévu théoriquement pour le HR
- Ménage scratch DSI : mail à faire à adminhpc pour éviter les suppressions de simulations mgie (fait le 26/09)



Suite...

- Prochaine réunion à programmer