

Distribution des données et post-traitements

Post-traitement des données CMIP5 à l'IPSL

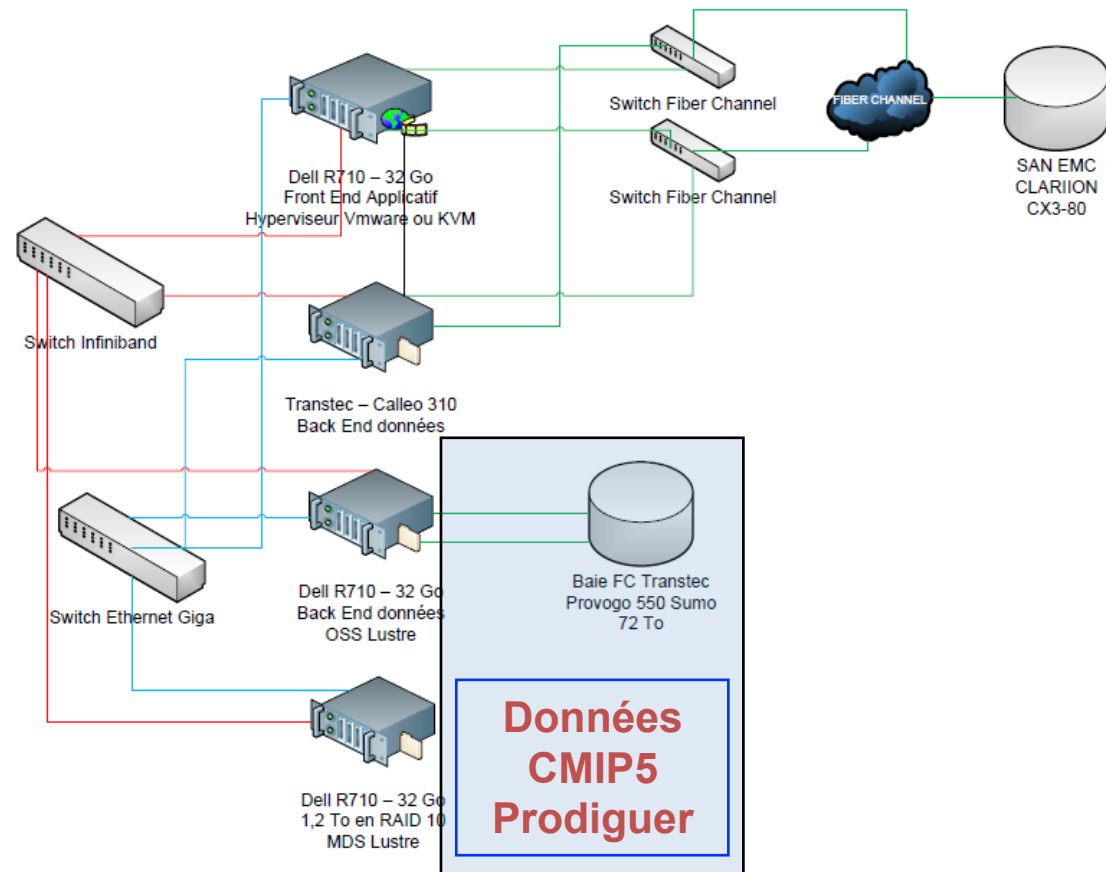
F. Corsini, S. Denvil , H. Moulron, K. Ramage, P. Weill

Infrastructure pour le post-traitement : Le mésocentre IPSL

❑ Plateforme Informatique IPSL Jussieu :

- Mise en place d'une infrastructure Infiniband (Réseau faible latence – 20 Gbit/s)
 - ✓ Entre les données et le calcul
 - ✓ Entre les serveurs de calcul

- Intégration de Prodiguer dans cette plateforme pour la distribution des données via le datanode ESG



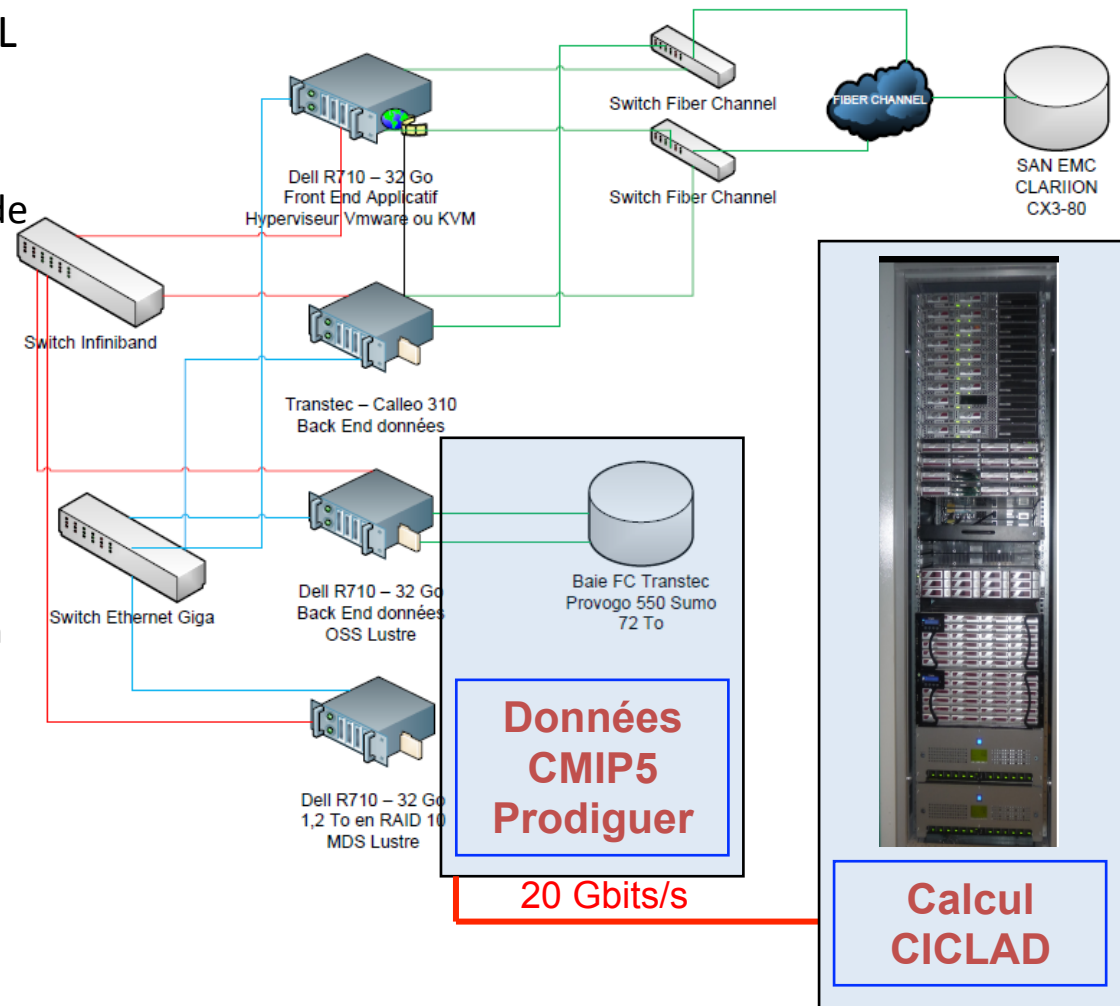
Infrastructure pour le post-traitement : Le mésocentre IPSL

□ CICLAD : Depuis septembre 2008, déploiement d'un cluster de calcul IPSL sur Jussieu, dans le cadre d'ESPRI, pour :

- les pré/post-traitements du pôle de modélisation
- L'exploitation des données Ether

➔ Pour CMIP5 :

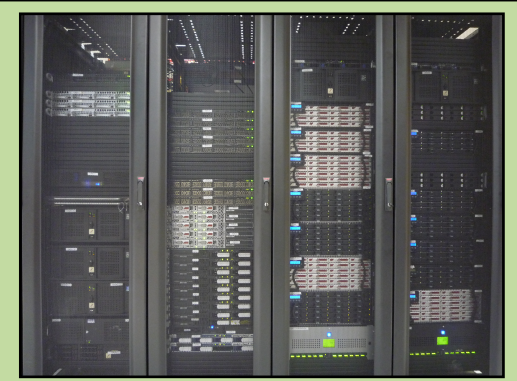
- Intégration de Ciclad à la plateforme IB
 - Accès aux données Prodiguer "en local" sur CICLAD
 - Inter-connexion des noeuds de calcul en IB
- Moyens pécfiques (stockage, calcul) mis en place autour de CICLAD



Infrastructure pour le post-traitement : Le mésocentre IPSL

❑ Lien vers les données d'observation disponibles sur le serveur de données ClimServ :

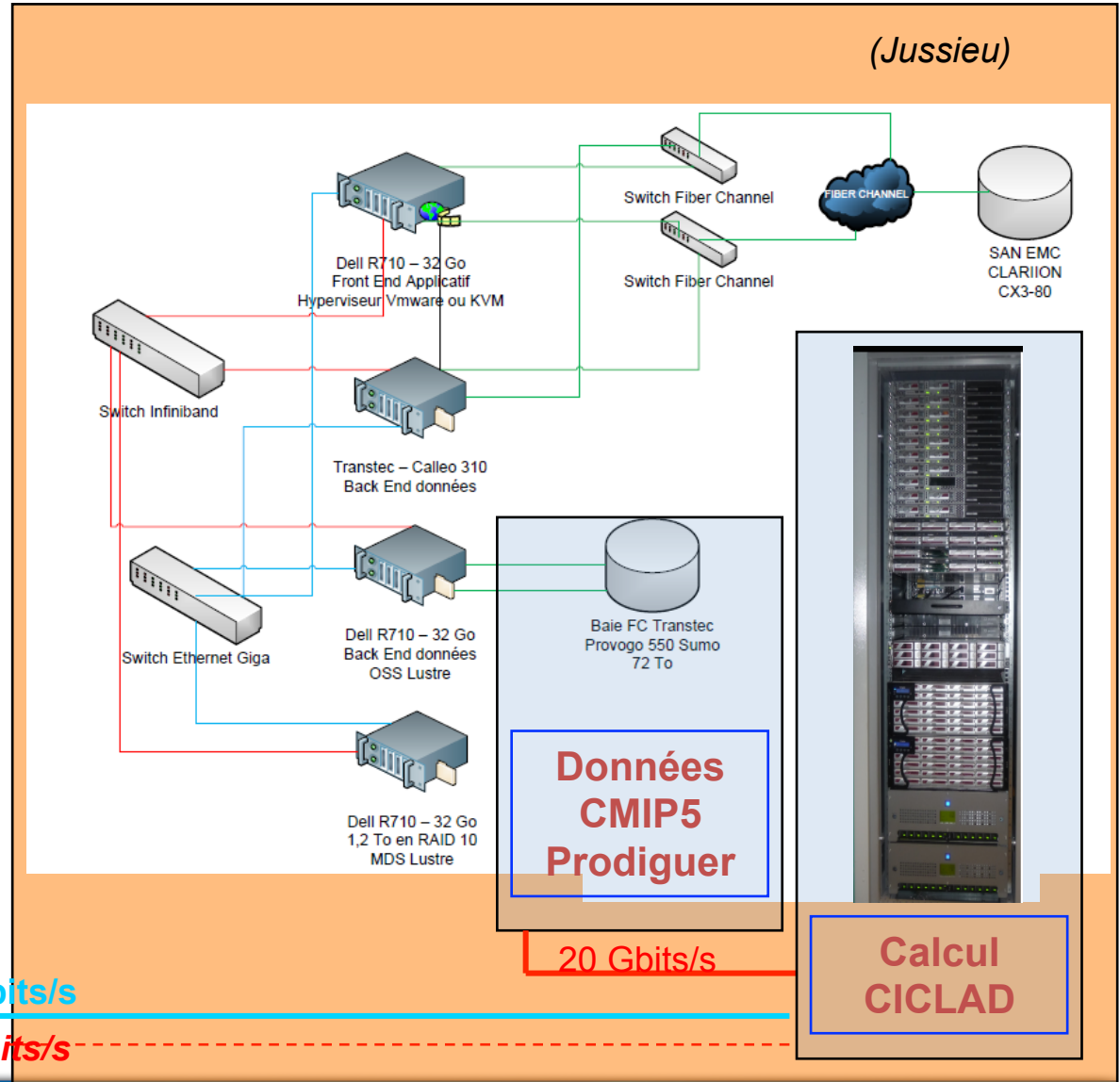
- FTP
- OpenDAP
- Accès local sur Ciclad (en projet)



**Données Observation
 ClimServ**

(Palaiseau)

100 Mbits/s
 - 10 Gbits/s



Infrastructure pour le post-traitement :

CICLAD (Calcul Intensif pour le Climat, l'Atmosphère et la Dynamique)

❑ Documentation et **demande de compte** :

<http://ciclad-web.ipsl.jussieu.fr>

❑ **Administrateur** : Philippe Weill (Philippe.Weill@latmos.ipsl.fr)

❑ **Calcul** :

- 10 noeuds de calcul AMD Dual quad-core 2,3GHz 32Go RAM
- 2 noeuds de calcul AMD Six-Core 2,6 GHz 48 Go RAM
- 96 coeurs accessibles en batch + 16 coeurs interactifs depuis les frontales Ciclad1 et Ciclad2
- *24 à 48 coeurs supplémentaires pour le traitement CMIP5*

❑ **Stockage** :

- /home : limité à 5 Go par utilisateur (sauvegardé)
- /data : 80 To pour d'espace de travail utilisateurs (non sauvegardé)
- /prodiguer : *≈ 120 To sous-jeux simulations CMIP5 (accessibles en lecture seulement)*
- /cmip5_wkspc : *≈ 10-30 To pour l'exploitation des données CMIP5 ou le rapatriement de données non-présentes dans Prodiguer*



Infrastructure pour le post-traitement : Logiciels de traitement

❑ Les logiciels de traitement disponibles :

- Bibliothèques :
NetCDF4, HDF5
- Traitement des données :
nco, cdo
- Analyse, visualisation :
Ferret, Matlab, Python, CDAT, R, ncview, GrADS, NCL
- Compilateurs (Fortran, C) :
PGI, Intel, NAG, GNU
- Parallélisme :
OpenMP, lammpi, openmpi
- Bibliothèques suggérées : **Statpack** (Opérations d'analyse statistiques en entrées/sorties NetCDF)

- **CFMIP Observation Simulator Package (COSP)**

❑ Mise en communs d'outils d'analyse

- Changement de grille (interpolation)
- calcul de moyennes saisonnières
- Valorisation des outils développés pour les précédentes analyses (Wiki, Repository...)

Données pour le post-traitement : Jeux d'observation de validation

☐ Les jeux d'observation pour la validation (liste établie à partir de l'enquête CMIP5) :

- Précipitation :
 - ✓ Satellite : CMAP, GPCP, TMI, SSMI
 - ✓ Sol : GPCC

- Océan :
 - ✓ SST TMI, HadISST

- Nuages :
 - ✓ CFMIP Observation (Cloudsat, Calipso, Parosol, AIRS)

- Bilan Radiatif :
 - ✓ CERES, ERBE
 - ✓ OLR

- Analyses, Réanalyses :
 - ✓ ERA, NCC (NCEP corrected by CRU), CRU, NCEP

Données déjà disponibles

**Données pouvant être rapatriées et mises en commun
sur les serveurs du centre de données IPSL**