

# Le couplage océan-atmosphère parallèle: d'OASIS3 à OASIS4

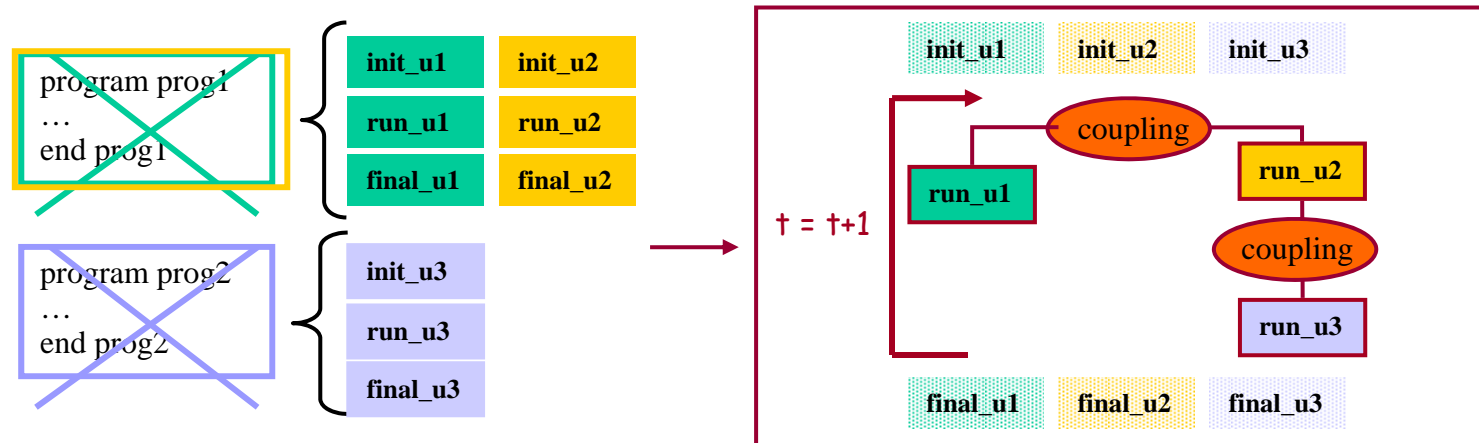
*S. Valcke, E. Maisonave, L. Coquart  
CERFACS*

- Quelques conclusions du workshop "Coupling Technologies for Earth System Modelling : Today and Tomorrow"
- OASIS3 & OASIS4:
  - Un survol historique
  - Communication
  - Interpolations et transformations
  - Quelques chiffres
- Conclusions

- CERFACS, Toulouse, 15-17 décembre 2010
- 45 participants:
  - France: 20
  - US: 11
  - EU: 12
  - China: 2
- 8 présentations sur les coupleurs et infrastructures de couplage existants:
  - US: ESMF (DoD, NASA, NSF, NOAA), CESM-CPL7 (NCAR), FMS (GFDL)
  - EU: OASIS (FR), BFG (UK), OpenMI (NL), PALM (FR), OOPs (ECMWF)
- D'autres présentations et beaucoup de discussions (voir <https://is.enes.org>)
- Une conclusion: il existe deux approches principales par rapport au couplage de codes:
  - intégrée, à-la-ESMF
  - à-la-OASIS

## Approche intégrée: infrastructure logicielle avec bibliothèque de fonctions

- Diviser les composantes originales en unités élémentaires init/run/finalize
- Adapter les structures de données et l'interface d'appel de ces unités
- Ecrire ou utiliser des unités de couplage (transformation, interpolation)
- Utiliser l'infrastructure logicielle pour bâtir un code unique intégrant les composantes

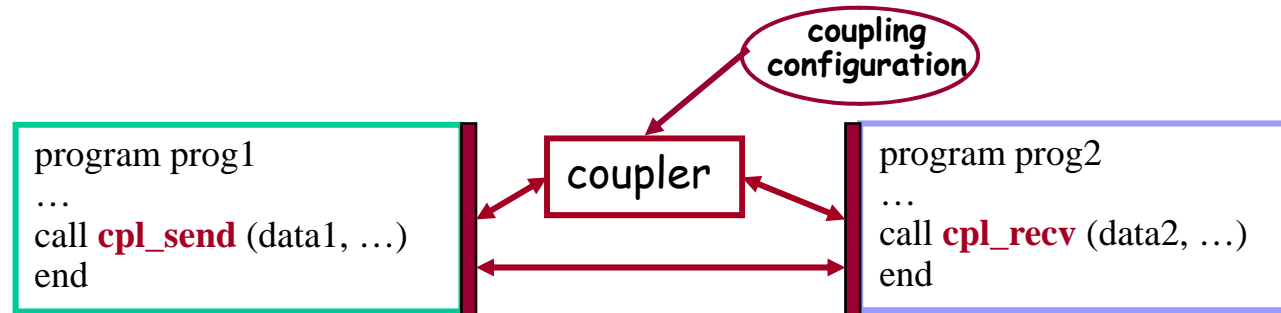


- ☺ Flexible
- ☺ Efficace (partage de la mémoire)
- ☺ Composantes séquentielles et concurrentes

- ☹ codes existants
- ☹ facile

Ex: ESMF, FMS, CPL7

## Approche « a-la-OASIS »: un coupleur et une librairie de communication

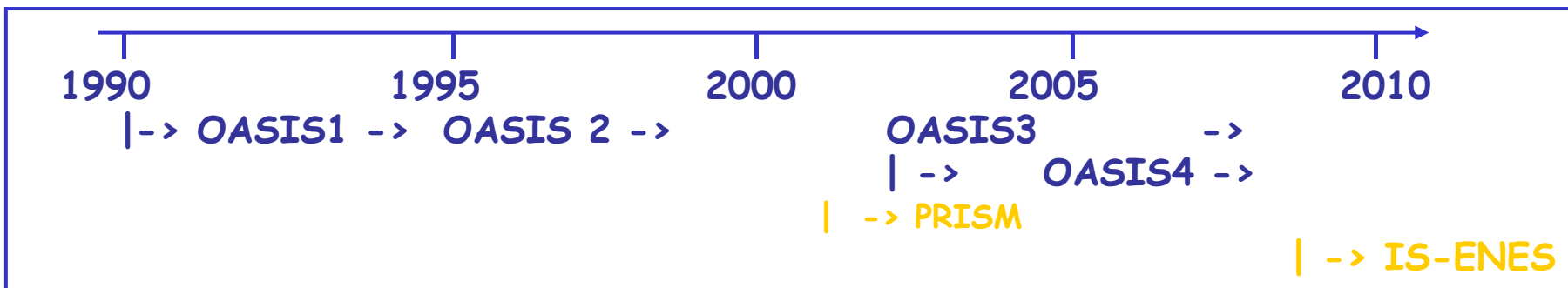


- ☺ Codes existants
- ☺ Composantes concurrentes

- ☹ Composantes sequentielles
- ☹ (efficient)

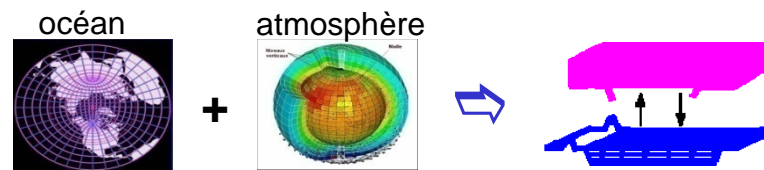
Ex: PALM, OASIS

- **L'outil de couplage idéal devrait combiner les deux approches** car l'approche adaptée dépend de la possibilité/volonté de changer les codes originaux, l'environnement de calcul, la séquentialité/concurrence des codes originaux, l'efficacité/performance recherchée, ...
- Est-on prêts en Europe à accepter des contraintes fortes de codages pour bénéficier des avantages de l'approche intégrée?
- Est-ce que l'évolution des machines/OS nous forcera à le faire?



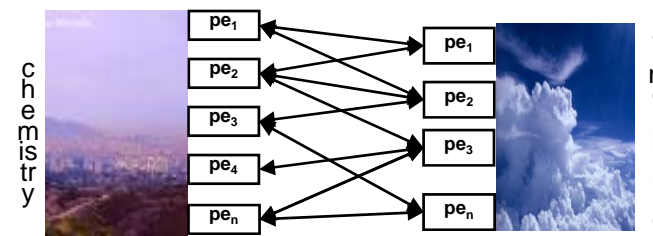
• **OASIS1 -> OASIS2 -> OASIS3:**

Couplage 2D océan-atmosphère  
 basse résolution, faibles fréquences de couplage  
 → flexibilité, modularité, interpolations 2D



• **OASIS4:**

Couplage 2D/3D de composantes massivement parallèles  
 sur des plateformes massivement parallèles  
 → parallélisme, efficacité, performance, interpolations 3D



- F90 et C
- logiciel libre distribué sous une licence LGPL
- Toutes les bibliothèques externes sont du domaine publique (MPI, NetCDF, XML) ou libres (LANL SCRIP, GFDL mpp\_io)
- Les développeurs actuels: CERFACS, CNRS, DKRZ, MPI-M
- ~30 groupes utilisateurs pour OASIS3 , ~5 pour OASIS4



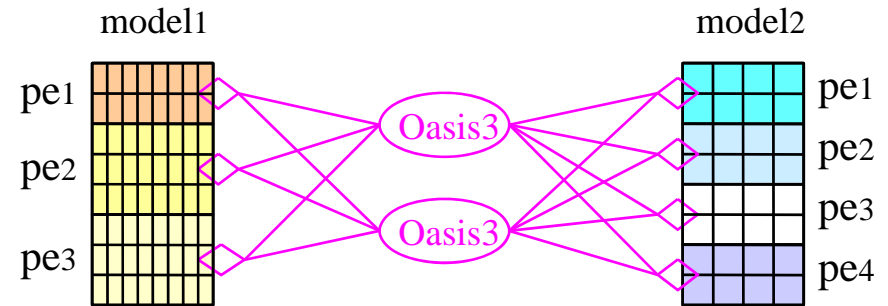
- Les composantes modèles restent des exécutables séparés
- OASIS =
  - une librairie de communication (PSMILe) liée aux composantes +
  - un autre exécutable (interpolation)

## Pour utiliser OASIS:

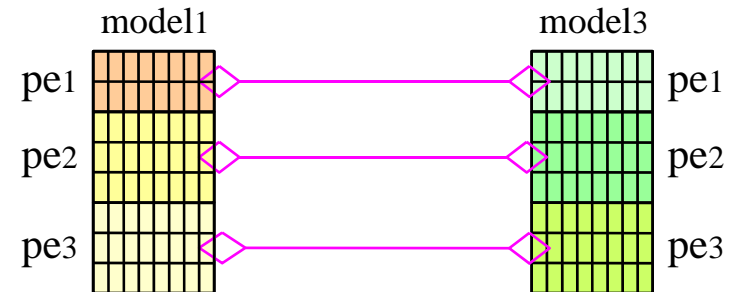
- Identifier les composantes modèles et les champs de couplage
- Interfacer les composantes modèles avec la librairie PSMILe:
  - call prism\_init
  - call prism\_def\_grid
  - call prism\_put (... , date, var\_array, ...)      call prism\_get (... , date, var\_array, ...)
- Choisir les paramètres de couplage (source et cible, fréquence, interpolations, etc.) et écrire le fichier de configuration
- Compiler OASIS et les composantes modèles liées à la PSMILe
- Démarrer les modèles et OASIS et le laisser gérer le couplage!

## PSMILe basée sur MPI1 or MPI2

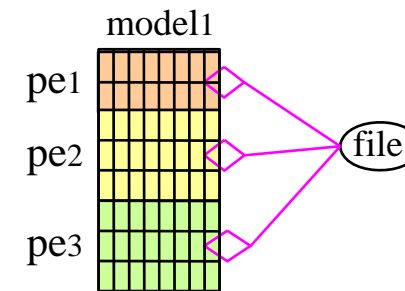
- Communication partiellement parallèle entre les composantes parallèles et les processus OASIS3



- Communication directe (même grille, même partitionnement)

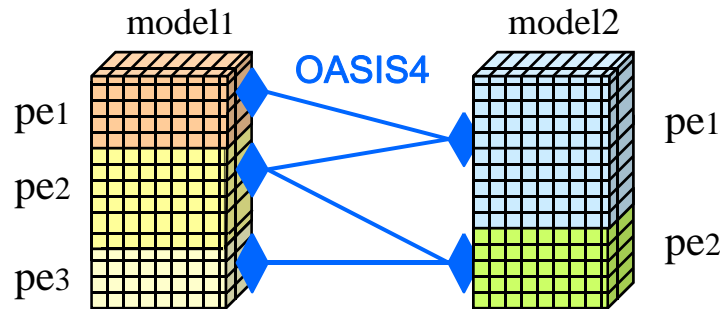


- I/O (bibliothèque mpp\_io du GFDL)

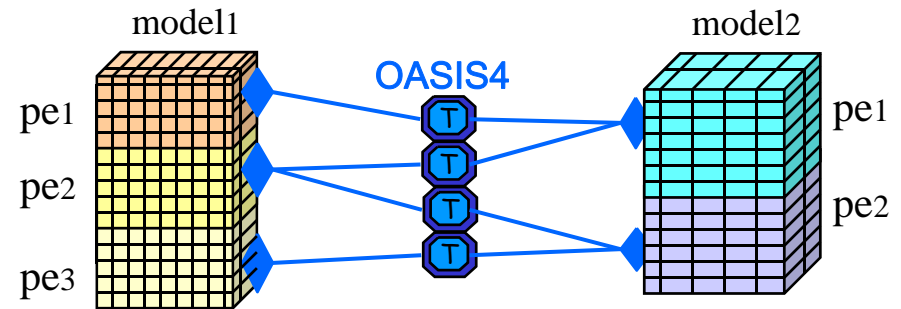


**Communication complètement parallèle, déterminée par la recherche des voisins pour l'interpolation faite par la PSMILe du coté source:**

- extraction de la partie utile du champ de couplage

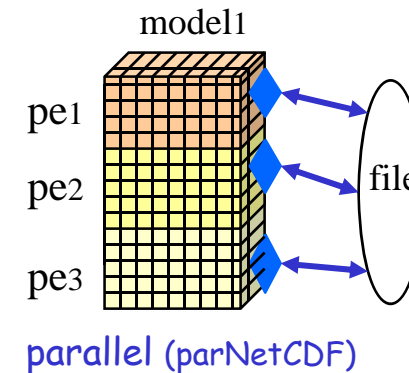
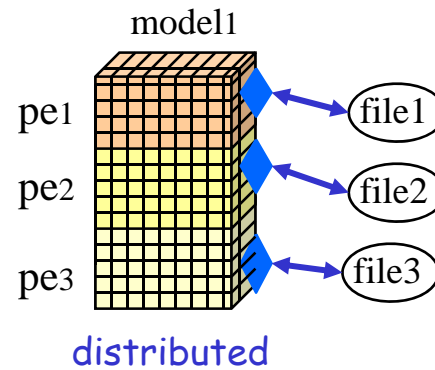
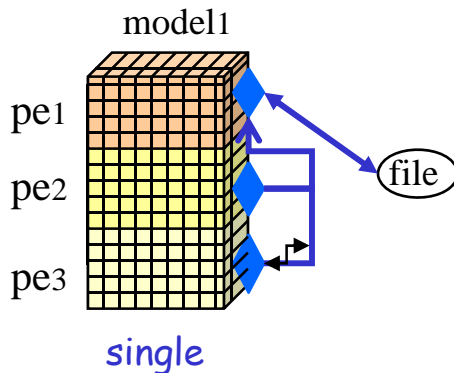


**Mêmes grilles, différents partitionnements**  
 ➤ repartitionnement direct



**Grilles et partitionnement différentes**  
 ➤ remaillage par OASIS4 parallèle

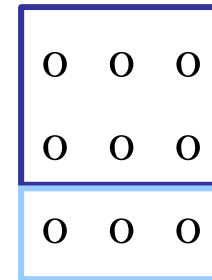
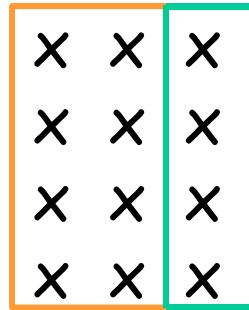
**I/O parallèles: GFDL mpp\_io + parNetCDF**





Exemple:  
1 + proche-voisin

Source  
2 processus



Cible  
2 processus



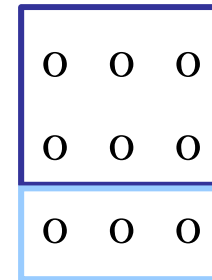
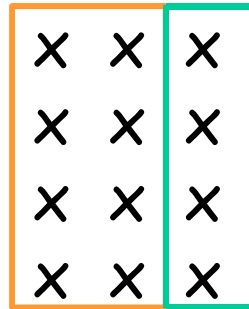
Pour chaque processus cible: (e.g. processus cible 1):

1/ Echange des enveloppes	2/ Recherche des voisins	3/ Def des "subsets" associés
	<p>source PE1</p>	
	<p>source PE2</p>	

➤ Le Transformer OASIS4 traite en parallèle les différents sous-ensembles de points sources et cibles associés pour chaque intersection de processus source et cible.

Exemple:  
2+proches  
- voisins

Source  
2 processus



Cible  
2 processus

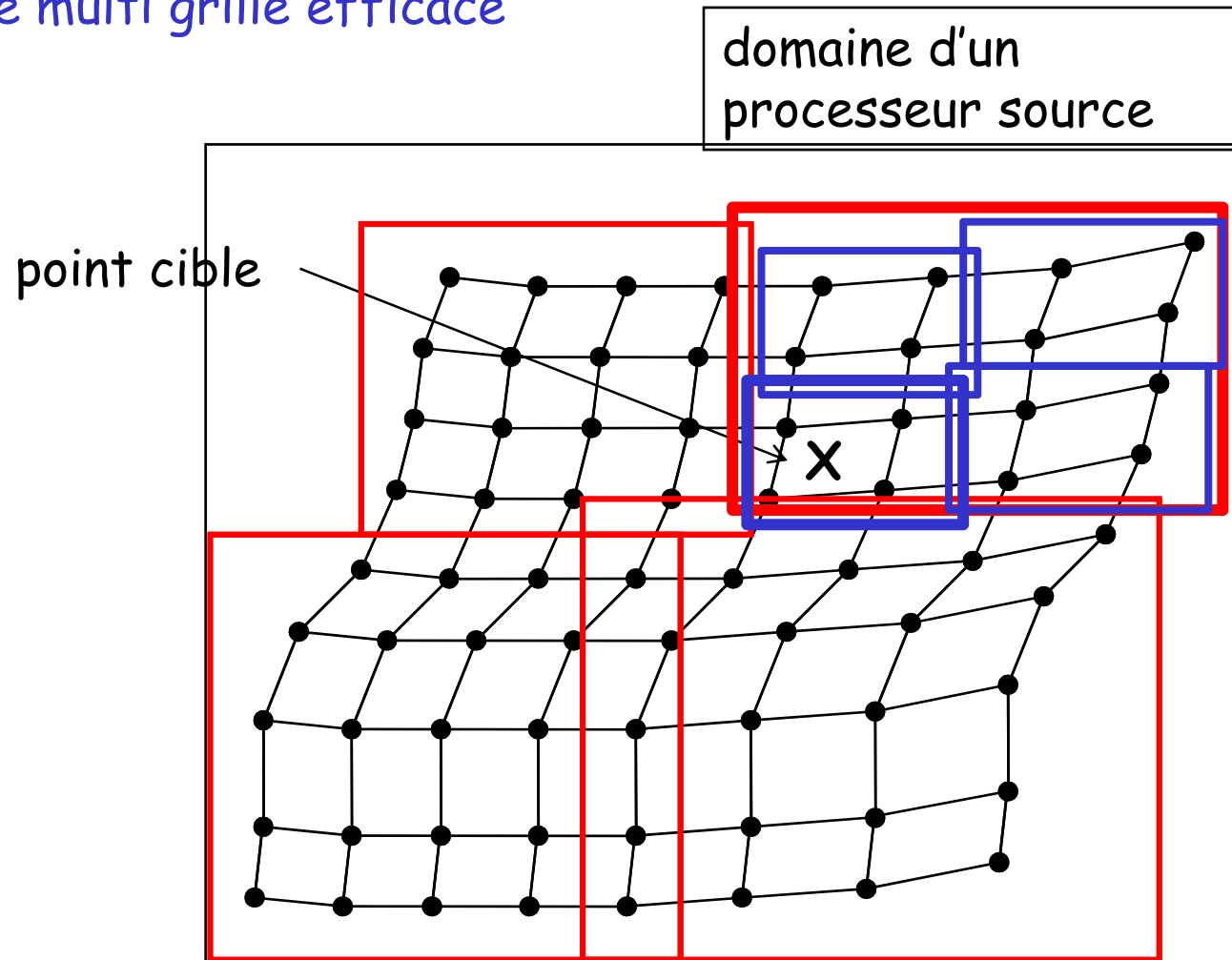


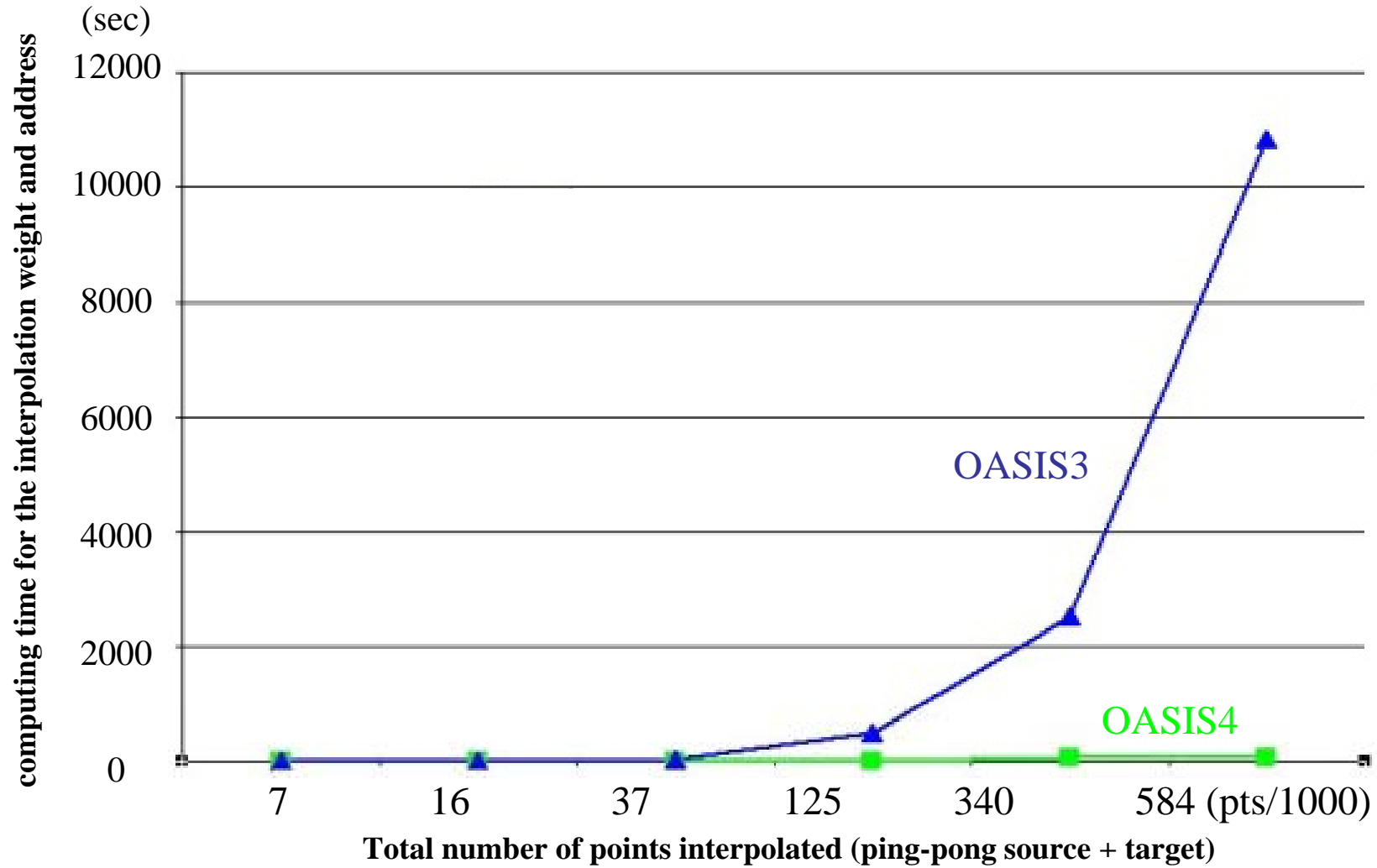
Pour chaque processus cible: (e.g. processus cible 1):

1/ Echange des enveloppes	2/ Recherche des voisins	3/ Def des "subsets" associés
	<p>source PE1</p>	<p>subsetS1</p> <p>subsetT1</p>
	<p>source PE2</p>	<p>subsetS2</p> <p>subsetT2</p>

➤ Le Transformer OASIS4 traite en parallèle les différents sous-ensembles de points sources et cibles associés pour chaque intersection de processus source et cible.

- Pour chaque point cible, la recherche des voisins se fait selon un algorithme multi grille efficace





## Quelques "gros" modèles couplés utilisant OASIS3:

- NEMO - LMDz (IPSL)
  - SGI ALTIX ICE (CINES)
  - LMDz ~1/3 deg, 589 Kpts, 39 niveaux verticaux
  - NEMO ORCA0.25, 1.5 Mpts, 75 niveaux verticaux
  - Couplage toutes les 2 heures
    - Avec 2191 cpus: 2048 LMDz, 120 NEMO, 23 OASIS3
- NEMO/CICE - UM (UK Met Office / Hadley Centre)
  - IBM power6
  - UM atm ~ 2/3 deg, 432x325 = 140 000 pts, 85 niveaux verticaux
  - NEMO/CICE ORCA0.25 : 1.5 Mpts, 75 niveaux verticaux
  - Couplage toutes les 3 heures
    - Avec 288 cpus: 192 UM, 88 NEMO, 8 OASIS3 : <2% de surcout de couplage

## Quelques "gros" modèles couplés utilisant OASIS3:

### •EC-Earth

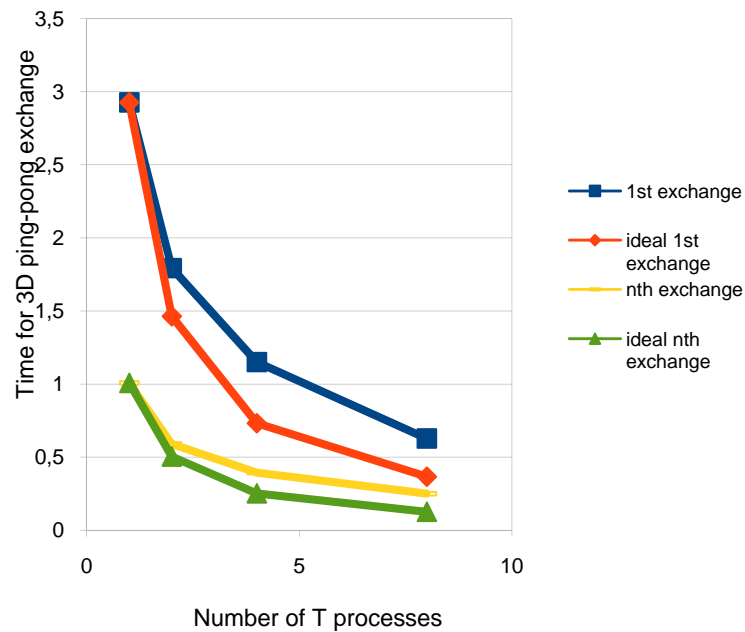
- Ekman cluster (Stockholm) : 1268 noeuds (chacun avec 2 quadripro AMD Opteron)
- IFS T799, 843 Kpts, 62 niveaux verticaux
- NEMO ORCA0.25, 1.5 Mpts, 45 niveaux verticaux

IFS-NEMO-OASIS nb of cores	512-128-1	512-128-10	800-256-1	800-256-10
1-IFS standalone	41.	41.	29.9	29.9
2-IFS coupled	45.7	n/a	32.7	n/a
3-NEMO coupled	38.5	n/a	24,6	n/a
4-EC-Earth3	45.7	42.3	33.2	30.3
5-OASIS	5.5	n/a	6	n/a
6-Coupling overhead	4.7 (13.4%)	1.3 (3%)	3.3 (11%)	0.4 (1.3%)

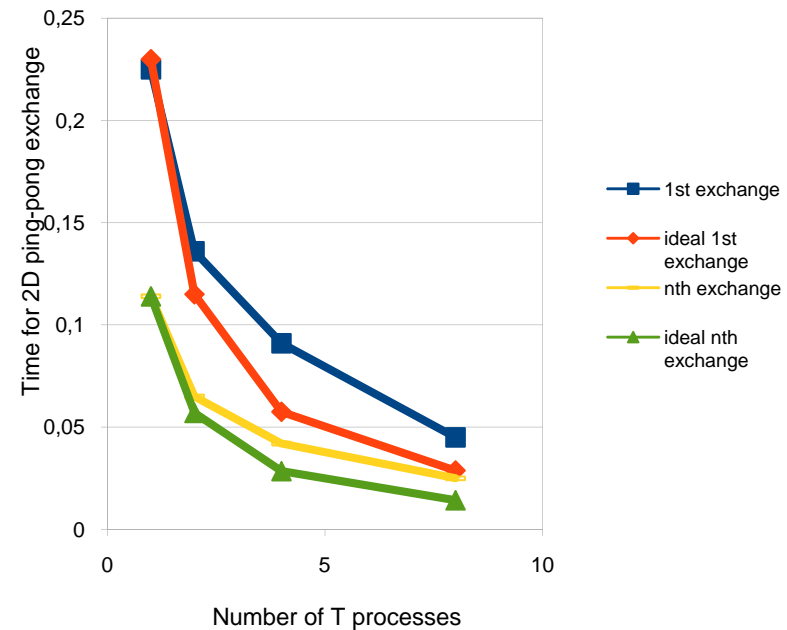
## Premiers tests de scalabilités:

- Modèle couplé jouet avec échanges ping-pong
- Composante "atmosphérique" 768× 385 et 40 niveaux verticaux (3D)
- Composante "océanique" 1202× 665 et 45 niveaux verticaux (3D)
- Cluster de PC local, compilateur Intel Fortran v10.1, GNU C v4.1.2, MPICH-GM

Transformer scalability (3D exchanges)

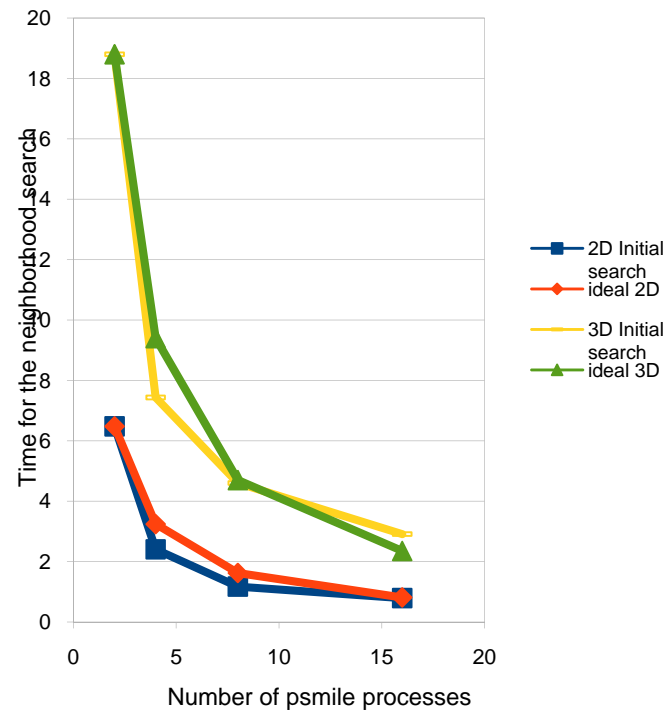


Transformer scalability (2D exchanges)



## Premiers tests de scalabilités:

Psmile scalability





- GEMS /MAC (ECMWF, Meteo-France, and KNMI):
  - couplage 3D dynamique atmosphérique - chimie atmosphérique
- B of Meteorology, Melbourne, Australia
  - couplage 2D entre MOM4p1 et UM6.4 (modèle à aire limitée - cyclone tropical Ului)
  - Plateforme: SUN constellation - 576 nœuds 2 quad-core Intel 64-bit Xeon Nehalem
  - MOM4p1 (110 Kpts x 47 lev) 24 cœurs, UM6.4 (90 Kpts x 50 lev) 160 cœurs, OASIS4 16 cœurs
  - couplage à chaque pas de temps
- Alfred Wegener Institute, (Bremerhaven, Germany)
  - couplage 2D océan-atmosphère entre FESOM and ECHAM5
- SMHI:
  - couplage régional 2D océan-atmosphère (RCO-RCA)

En cours de développement:

- ARPEGE (T359) - NEMO 0.25 , Météo-France NEC SX9 (CERFACS)
- ARPEGE (T359) - NEMO 0.25 , CINES SGI Altix ICE (CERFACS)
- ECHAM6 T159 - MPI-OM 0.4 deg au DKRZ (MPI-M)

## OASIS3

- Stable, bien débogué , large communauté, vraie "success story"!
- Design simple, parallélisme limité, n'importe qui peut déboguer !
- Support utilisateur actif, pas de nouveaux développements (ou presque)

## OASIS4

- Interpolation et communication complètement parallèles et efficaces
- Encore beaucoup de débogage en cours ☹
- Plus complexe, plus fragile, seul un expert peut déboguer
- Adoption lente par la communauté
- Limitation principale : pas de grilles non- structurées

## Et après?

- Développement et support à plus long-terme (MoU?) en cours de discussion dans IS-ENES
- Evolution vers d'autres outils (PALM?), vers l'approche intégrée?

Fin, ouf !