INSTITUT PIERRE SIMON LAPLACE
DES SCIENCES DE L'ENVIRONNEMENT

Stratégie calcul en modélisation du climat

Mars 2007

Document technique

Cette note présente l'état d'avancement de la modélisation du climat et ses besoins en moyens de calcul intensif sur les machines du CNRS et du CEA pour les années à venir, dans le contexte de la préparation du prochain rapport du GIEC. Elle fait le point sur les capacités présentes et futures d'utilisation des différentes architectures de machine. Elle est constituée d'une synthèse et de ce document technique détaillé.

Cette analyse s'appuie sur le Pôle de modélisation de l'IPSL rassemblant 80 chercheurs et 20 ingénieurs (CEA, CNRS, Ecole Polytechnique, ENS, IRD, UPMC, UVSQ, …). L'autre groupe Français de modélisation climatique à Météo-France dispose de moyens de calcul dédiés.

Le document se présente sous la forme de 3 parties autonomes présentant :

1. Le plan de travail pour les prochaines simulations GIEC[1] prévue dès 2009 ;
2. L'état actuel de la parallélisation des codes et du modèle couplé de l'IPSL ;
3. Les verrous à lever pour que les modèles de climat utilisent de façon efficace les machines massivement parallèle/scalaire.

Document rédigé par le Pôle de modélisation de l'IPSL

[1] GIEC : Groupe Intergouvernemental d'Expert sur l'Évolution du Climat (en anglais, IPCC : International Panel on Climate Change), http://www.ipcc.ch.

I. Plan de travail pour le GIEC 2009

I.1. Introduction

Notre capacité à prévoir les modifications climatiques majeures susceptibles de se produire au cours des prochaines années ou des prochains siècles est fortement conditionnée par l'intégration dans des ensembles complexes de processus et de couplages ayant des échelles de temps et d'espace très variées. Les simulations imposées par le GIEC[2] permettent d'obtenir des ensembles de simulations climatiques, prenant en compte les incertitudes des modèles actuels, à partir desquels il est possible d'établir l'état le plus probable du système climatique. Elles servent également de référence à de nombreuses études.

Les projections climatiques dont les résultats ont servi de référence pour le 4ème rapport du GIEC (AR4[3]) étaient au nombre de 11. La version du modèle de l'IPSL utilisée (IPSLCM4) avait été figée en juillet 2004 après 4 ans de développement. Les simulations ont représenté un total de 80 000 heures de NEC SX-5 sur les deux centres de calculs CCRT[4] (CEA) et IDRIS[5] (CNRS), soit près de 20 processeurs NEC (SX-5 et SX-6) simultanément pendant la phase de production intensive (6 mois). L'exercice s'est déroulé dans une période de forte pénurie de moyens de calcul en France, soit 3 à 7 moins que nos partenaires européens (voir le rapport de l'Académie de technologies[6] et le rapport Héon et Sartorius[7]). L'adéquation entre moyens de calcul et simulations à réaliser s'est traduite par des simulations ayant la plus basse résolution de l'ensemble multi-modèles. Chaque simulation de 300 ans était réalisée en 2 à 3 mois de temps réel. Cette unité de 2 mois de temps réel est une constante que l'on retrouve dans les différentes équipes internationales ayant participé à l'exercice du GIEC, chacun ayant adapté la résolution aux ressources disponibles (voir l'annexe V.2).

Nous préparons dès à présent les modèles qui permettront d'aborder les nouvelles questions pour le prochain exercice du GIEC en vue du cinquième rapport (AR5). Ce travail allie l'exploitation des simulations existantes, la réalisation de simulations complémentaires, le développement de nouvelles paramétrisations physiques, et les développements informatiques permettant d'améliorer le contenu des modèles ou leurs performances.

I.2. Plan de route pour le prochain exercice GIEC (2009)

I.2.1. Les enjeux scientifiques

Progresser dans la compréhension et la prévision des changements climatiques demande d'augmenter nos systèmes et de développer la modélisation sur plusieurs fronts. Il faut tout à la fois améliorer le contenu physique des modèles, mieux représenter les phénomènes de petite échelle, comprendre les interactions complexes entre le climat et les cycles biogéochimiques. Il faut aussi augmenter la complexité et les couplages entre les processus, étudier les incertitudes liées à la nature chaotique du système en réalisant des ensembles de simulations, et tester la possibilité d'évènements peu probables, mais à haut risque pour la société. Ces

---

[2] GIEC : Groupe Intergouvernemental d'Expert sur l'Évolution du Climat (en anglais, IPCC : International Panel on Climate Change), http://www.ipcc.ch.
[3] AR5 : 5ème rapport d'évaluation (assessment report) du GIEC. Prévu pour 2012. Le 4ème rapport (AR4) est paru en février 2007.
[4] CCRT : Centre de Calcul de la Recherche et la Technologie, Bruyères-le-Châtel.
[5] IDRIS : Institut du Développement et des Ressources en Informatique Scientifique, Orsay.
[6] http://www.academie-sciences.fr/comites/pdf/groupe_ics_06_12_06.pdf.
[7] http://www.recherche.gouv.fr/rapport/index.htm#calcul.

grands enjeux ont été déclinés dans différents documents de prospective, incluant les réflexions sur le calcul intensif au niveau européen (document HPC Europe).

Ces évolutions vont toutes accroître les ressources nécessaires (calcul, mémoire, espace fichiers, outils d'analyse, …). Elles se caractérisent par :

- la variété des simulations pour pouvoir prendre en compte le système dans son ensemble ou le rôle de l'une ou l'autre composante ou paramétrisation ;
- l'augmentation de la résolution horizontale et verticale ;
- l'augmentation des longueurs des simulations ;
- la généralisation des simulations d'ensemble et l'augmentation du nombre de simulations dans chaque ensemble ;
- la complexification des composantes (plus d'espèces chimiques, plus de processus physiques, …) ;
- des couplages de plus en plus poussés et offrant des possibilités d'analyses croisées avec nos partenaires.

I.2.2. Les versions des modèles

Pour remplir les différents enjeux scientifiques nous développons deux versions du modèle couplé de l'IPSL. La première permettra d'augmenter la complexité en prenant en compte les cycles biogéochimiques, alors que la deuxième sera destinée à mieux représenter les différents éléments du cycle diurne et les phénomènes climatiques de petite échelle.

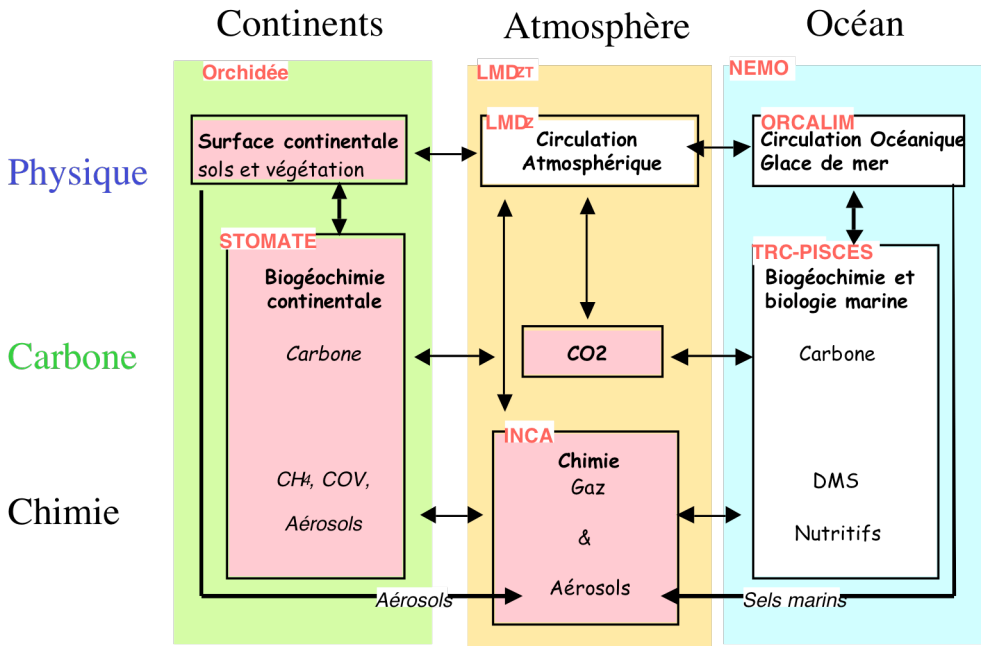

Figure 1 : le modèle système climatique de l'IPSL.

Le modèle océan-atmosphère actuel IPSLCM4 (Figure 1) est composé des modèles de circulation générale atmosphérique (LMDZ[8]) et océanique (NEMO[9], ex-OPA), dialoguant au

[8] http://www.lmd.jussieu.fr/~lmdz/
[9] http://www.lodyc.jussieu.fr/NEMO.

travers du coupleur OASIS[10], chargé des interpolations entre grilles, des conversions d'unités et des changements de convention. La composante atmosphérique LMDZ inclut la composante continentale ORCHIDEE[11]. De la même façon, la composante océanique NEMO inclut la composante glace de mer (LIM). Le modèle climat-carbone IPSLCM4_LOOP est basé sur le même modèle IPSLCM4 auquel ont été ajoutés les processus liés au carbone continental (ORCHIDEE/STOMATE), à son transport dans l'atmosphère (LMDZT) et à son absorption dans l'océan (NEMO/PISCES). Ces deux modèles ont été utilisés pour les simulations du GIEC réalisées en 2005 et 2006. Développé en 2006 en cohérence avec le modèle climat IPSL, le modèle incluant les aérosols et la chimie intègre de plus un module de chimie de l'atmosphère (INCA[12]), plus ou moins complexe selon la chimie utilisée. Cette version du modèle couplé commence à être utilisée pour des études scientifiques.

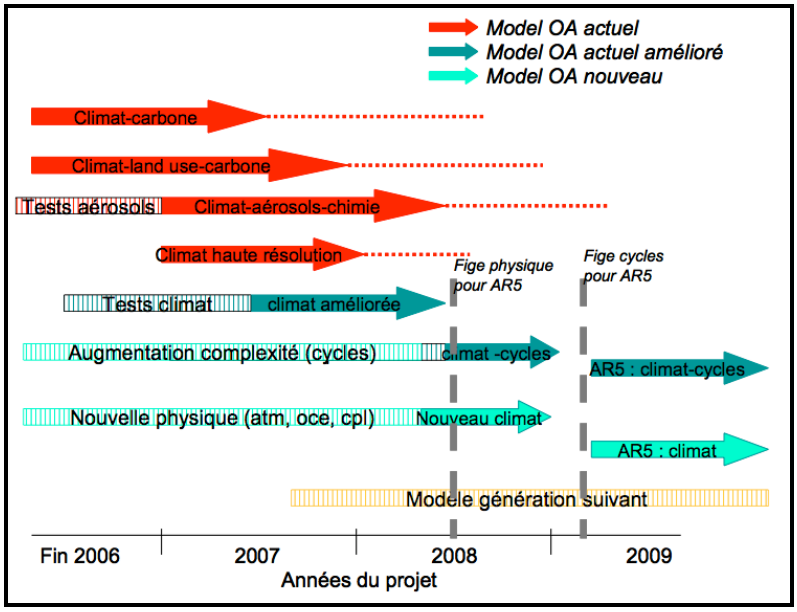

Figure 2 : Plan de route du pôle de modélisation du climat de l'IPSL pour la réalisation des prochaines simulations du GIEC prévues en 2009. Ce plan souligne les principales chaînes couplées que nous devons mener en parallèle, les nouveaux développements et les dates clefs permettant de figer les versions de modèles qui seront utilisées pour l'AR5. Cette action est soutenue par le projet MISSTERRE de l'INSU (LEFE).

Le Tableau 1, présente les différents modèles utilisés à l'IPSL, seuls ou couplés à une autre composante. A partir des modèles et des couplages existants plusieurs étapes sont prévues. La première consiste à corriger des biais identifiés dans la version IPSLCM4 utilisée pour l'AR4. Cette nouvelle version de la physique du modèle devra être figée début 2008 pour se donner le temps de valider les différentes versions incluant les cycles biogéochimiques (Figure 2). Les modèles biogéochimiques bénéficieront de nouveaux développements, comme le cycle de l'azote ou du méthane. Pour l'AR4, la résolution utilisée était de 3,75°x2,5°x19 niveaux pour l'atmosphère (96x71x19) et 2°x1,5°x31 niveaux pour l'océan (modèle ORCA2 182x149x31). Nous continuerons à utiliser cette résolution pour les couplages avec les cycles biogéochimiques.

Les nouveaux développements physiques seront incorporés dans une nouvelle version du modèle couplé (Figure 2). Ils permettront une meilleure prise en compte du cycle diurne, en lien avec le cycle de vie des nuages et de la convection, des échanges plus fréquents entre l'océan et l'atmosphère, et une meilleure représentation des processus physiques sur la verticale. Pour résoudre les trajectoires des dépressions des moyennes latitudes (*storm tracks*), la résolution atmosphérique idéale est de 1°x1°x50 niveaux (360x180x50). Son coût est tel que nous réali-

[10] http://www.cerfacs.fr/globc/software/oasis/index.html
[11] http://orchidee.ipsl.jussieu.fr.
[12] http://www-lsceinca.cea.fr.

serons cependant une grande partie des expériences à une résolution de l'ordre de 2,5°x1,25°x50 niveaux (144x142x50). Le modèle d'océan sera ORCA05, qui a une résolution de 0,5°x0,5°x48 niveaux (720x560x48).

Laboratoire(s) développeur(s)	IPSL	Non IPSL	International
LMDZ	LMD, LSCE, IPSL	LGGE, LOA	
ORCHIDEE	LMD, LSCE, IPSL		
INCA	LSCE, LMD, SA		
NEMO-OPA	LOCEAN	LEGI, LMC, LPO, Mercator-ocean	UK Met Office, SOC
NEMO-TOP	LOCEAN, LSCE		
NEMO-LIM	LOCEAN		UCL/ASTR
OASIS		CERFACS	
Système couplé	LMD, LSCE, IPSL	LGGE	
IOIPSL	LSCE, IPSL		
Modipsl, libIGCM	LSCE, IPSL		
Fast, Atlas	LSCE		

Tableau 1 : Tableau synthétique des composantes du modèle climat IPSL. Voir aussi le dossier demande de labellisation « Code numérique communautaire » et « Centres de traitement et d'archivage des données » pour la Plateforme de modélisation climatique de l'IPSL accessible à : http://dods.ipsl.jussieu.fr/mafoipsl/ESCI/Dossier_Label_IPSL.pdf.

De même, nous devons dès à présent préparer le modèle de climat de la génération suivante, ce qui constitue le dernier point du notre plan de route à 4 ans. Il deviendra envisageable avec ce type de modèle de travailler sur des maillages de 1000x1000, sujets au parallélisme massif. L'utilisation en standard n'est pas envisageable avant une dizaine d'années.

I.2.3. Le passage en parallèle pour tous les utilisateurs (de mi-2007 à mi-2008)

Un gros effort de parallélisation de nos codes a été entrepris ces dernières années. Nous développons une stratégie permettant d'être indépendant du type d'architecture de calculateur. Actuellement, nos codes peuvent tourner en parallèle lorsqu'ils sont utilisés seuls, y compris sur des machines scalaire/parallèle (cf. section II). Il reste pour l'atmosphère et les surfaces continentales à passer en pleine exploitation. Le modèle couplé peut tourner avec chaque composante en parallèle dans une version prototype. La mise en production sera effectuée en 2007 sur NEC SX8 (IDRIS et CCRT). Nous estimons qu'il faudra 1 à 1,5 années pour que cette version parallèle devienne l'outil standard de tous les utilisateurs. Cette période pourrait être raccourcie avec 4 personnes supplémentaires en mettant le portage et la maîtrise au quotidien du parallélisme en première priorité des scientifiques. Mais cette option impliquerait de repousser les développements en modélisation inscrits à notre plan de route (nouvelle physique, validation climatologie, cycle diurne, cycles biogéochimiques …).

Le projet ANR CICLE[13] nous a permis de compléter les équipes et le couplé parallèle est un des livrables du projet (mi 2007). Les développements et les premiers résultats de la version prototype montrent qu'aux résolutions utilisées actuellement (Tableau 2), 5 ou 6 processeurs seront dédiés à l'atmosphère et 1 processeur au coupleur et à l'océan. Ceci permet de rester sur un seul nœud NEC SX-8, de réduire les temps d'exécution et de ne pas trop surcharger les opérations de recombinaison de fichiers de résultats.

[13] projet ANR sur la ligne calcul intensif, 2006-2008. http://dods.ipsl.jussieu.fr/CICLE.

	Composantes	Résolution	Temps NEC SX8 (nb jours en machine pour 100 ans simulés)	Mémoire	Fichiers générés pour 100 ans
OA actuel IPSLCM4	LMDZ ORCHIDEE OASIS NEMO/OPA NEMO/LIM	**96x71x19** 3,75°x2,5°	5 ans/jour 20 jours/100 ans	4 Go	400 Go
		96x71**x50**	1,5 an/jour 66 jours/100 ans	8 Go	1 000 Go
		144x96x19 2,5°x 1,8°	2 ans/jour 50 jours/100 ans	6 Go	800 Go
		144x**142**x19 2,5°x1,27°	1 an/jour 100 jours/100 ans	8 Go	1 200 Go
		144x142x**50**	*Estimation* 360 jours/100 ans	*12 Go*	*1 800 Go*
		144x142x50 **ORCA05**	*Estimation* 400 jours/100 ans	*50 Go*	*7 000 Go*
		360x180x50 ORCA05	*Estimation* 2500 jours/100 ans	*160 Go*	*22 000 Go*
Carbone IPSLCM4_LOOP	LMDZT ORCHIDEE STOMATE OASIS NEMO/OPA NEMO/LIM NEMO/PISCES	96x71x19 ORCA2 **26 traceurs**	4 ans/jour 25 jours/100 ans	6 Go	1 000 Go
Chimie aérosols IPSLCM4_ESM	LMDZT ORCHIDEE STOMATE OASIS INCA NEMO/OPA NEMO/LIM NEMO/PISCES	89 traceurs 300 réactions	1 an/jour	9 Go	4 000 Go

Tableau 2 : coût des différentes configurations et résolutions utilisées actuellement.

L'exploitation du couplé sur machine scalaire/parallèle nécessite encore de prouver sa faisabilité, puis il s'agira d'optimiser ses performances. De plus les codes et leurs environnements doivent être facilement exploitables par des utilisateurs variées. En effet, la communauté concernée regroupe aussi bien des chercheurs, des ingénieurs que des stagiaires de master. Les tâches supplémentaires que nous avons déjà identifiées sont les suivantes :

- Installation optimale des outils de couplage (OASIS), de traitements de fichiers (bibliothèque NetCDF, nco, python), de distribution de fichiers (DODS/OpenDAP) ;
- Expertise sur les entrées/sorties et sur la recombinaison des fichiers. Éventuellement déploiement des solutions dans les centres de calculs.
- Benchmark de faisabilité : un an de simulation du modèle couplé associant les 3 exécutables, chacun sur un nombre différent de processeurs ;

- Benchmark de production : mesurant le temps de réalisation de 10 ans de simulation, incluant l'équilibrage des tâches sur les différents processeurs, la fourniture des résultats, la recombinaison des fichiers, la création des séries temporelles et des analyses ;
- Diffusion des bonnes méthodes de travail au sein des équipes, formation, support ponctuel ;

I.2.4. Besoins en calcul associés

Besoins sur les machines actuelles : NEC SX8

Le Tableau 2 résume les performances informatiques du modèle pour les configurations et les résolutions utilisées actuellement. Cette base permet d'établir les ressources en calcul nécessaires. Elles représentent 150 000 heures NEC SX-8 à l'IDRIS en 2007 et 120 000 heures NEC SX-8R au CCRT en 2007. Cela représente 25% des disponibilités à l'IDRIS et autant au CCRT. C'est également l'équivalent de 48 000 années simulées à la résolution 96x71x19 ou 19 200 années à la résolution 144x96x19. Cela uniquement pour l'activité directement concernée par les projections climatiques, des ressources équivalentes étant nécessaires au reste de l'activité de recherche en climat.

Les discussions en cours au niveau international et l'évaluation des différentes simulations à réaliser indiquent que le protocole d'expériences du GIEC s'achemine vers la réalisation d'ensembles de 6 simulations pour environ 5 scénarios, chacune pour 200 années de simulation soit un total de 6 000 années de simulation (Tableau 3). Pour cet exercice, une attention particulière sera portée au couplage avec le cycle du carbone et aux simulations à plus haute résolution. Pour l'IPSL, cela se traduit par un ensemble de simulations couplé climat carbone à la résolution atmosphérique 96x71x19 (soit 3,75°x2,5°) et océanique 182x149x31 (soit 2°x1,5°), un sous- échantillon à la même résolution prenant aussi en compte la chimie et les aérosols, un ensemble de simulation climatiques à la résolution atmosphérique 144x142x50 (soit 2.5°x1.27°x50) et du ½° pour l'océan (ORCA05), et quelques tests à haute résolution (1°x1°x50 x ORCA05). Avec les estimations des temps de calcul du Tableau 3, participer à l'AR5 représentera en 2009 un temps de calcul équivalent à 80 processeurs NEC SX-8.

Nos plans prévoient actuellement une utilisation conjointe des NEC SX-8 et SX-8R de l'IDRIS et du CCRT. Il est important de noter que la part des ressources calcul de l'IDRIS dédiées aux simulations GIEC n'a été que de 10 à 20% pour l'AR4, puisque entre 4 et 8 processeurs ont été réservés alors que le reste était normalement affecté à la production habituelle de l'IDRIS. Pour 2009 également, la production scientifique non directement reliée au GIEC doit pouvoir se réaliser dans des conditions non dégradées. La réalisation de ces simulations dans de bonnes conditions ne pourra se faire que si les moyens disponibles à l'IDRIS sont augmentés d'un facteur 2 ce qui correspond à l'arrivée de la deuxième tranche du SX 8 ou, à défaut, à l'utilisation pour la communauté climat de l'ensemble de la machine actuellement disponible. La moitié de la ressource sera employée pour les simulations exigées par le GIEC. L'autre moitié servira à des simulations de réglage, à des études de sensibilité non prévues par le GIEC, et à d'autres recherches en climatologie.

	GIEC AR4	***vers le GIEC AR5***
Tests et réglages de la physique, du climat, des métriques, des techniques d'analyses, d'atlas, …	Nombreux tests Entre août 2003 et juin 2004 **30 000 h NEC SX-5** soit l'équivalent de 2 500 années de simulation résolution de mise au point: 72x45x19/ORCA4 résolution retenue : 96x72x19/ORCA2	*2007-2008* *Ttests sur résolutions différentes* *Mise au point et utilisation généralisée des outils de lancement et d'analyse de simulations d'ensembles* ***5 000 années de simulation*** *ou 55 processeurs NEC SX-8 pendant un an*
Date de la version figée	Juillet 2004	*Mi 2008*
Simulations GIEC	Août 2004 - Juin 2005 **30 000 h NEC SX-5** soit 2 500 années de simulation - résolution retenue : 96x72x19/ORCA2	*Mi 2008-2009* ***IPSLCM4 LOOP (96x71x19)*** *5 scénarios de 200 ans, réalisés par des ensembles de 6 simulations chacun, soit un total de **6 000 ans** de simulation* ***IPSLCM4 ESM (96x71x19)*** *Environ **3000 ans** de simulation* ***IPSLCM4 (144x142x50)*** *Environ **6000 ans** de simulation* ***Résolution 360x180x50-ORCA05*** *Quelques scénarios selon les possibilités.* **80 processeurs NEC SX-8 pendant un an** *Utilisation simultanée du nombre de processeurs opportun pour avoir 150 ans simulés par mois*
Simulations GIEC climat carbone	Juin 2005 – Octobre 2005 60 000 h NEC SX-5 soit l'équivalent de 5 000 années de simulations	*Version figée début 2009* *Le double soit l'équivalent de 12 000 années de simulation*
Etudes de sensibilité variées	Juin 2005 – 2006 40 000 h NEC SX-5	*2009 et au delà.* **80 processeurs NEC SX-8 pendant un an**

Tableau 3 : tableau de marche et type de simulations à réaliser pour préparer le prochaine exercice du GIEC en 2009, suivant l'expertise acquise pour l'AR4.

Evaluation des besoins sur une machine scalaire/parallèle

Il est possible de chiffrer les besoins équivalents à la participation à l'AR5 sur machine scacaire/parallèle. Un rapide calcul permet d'estimer le nombre de processeurs scalaires

« équivalents » à nos besoins de calcul. En effet, une simulation à la résolution actuelle réalisable en 3 jours sur 8 processeurs de SX8 ne l'est qu'en 6 à 11 jours sur 16 processeurs scalaires (selon le type de processeurs). A cette résolution, l'utilisation d'une machine scalaire/parallèle est envisageable.

Nos besoins seraient alors équivalents à 800 à 1000 processeurs de type Itanium (le besoin en processeur de type Opteron reste à préciser). Néanmoins, comme nous l'avons précisé dans la section I.2.3, sur ces machines de nombreuses incertitudes sur les outils subsistent : environnement, les performances, temps de restitution des expériences complètes, … Les difficultés prévisibles concernent l'équilibrage du couplage, la gestion des entrées/sorties, le post-traitement en opérationnel, le traitement massif des fichiers, le passage en routine de simulations en couplé de centaines d'années, …

Il n'est donc pas raisonnable d'envisager l'utilisation de ce type de machine pour la réalisation des simulations de l'AR5.

I.3. Bilan des besoins pour l'AR5.

Dans la continuité de la prospective établie au moment du renouvellement de la machine vectorielle de l'IDRIS, nous avons besoin de l'augmentation en ressources prévues (facteur 10 à l'IDRIS en 2 tranches), qui correspond à l'arrivée de la deuxième tranche du SX 8 ou, à défaut, à l'utilisation pour la communauté climat de l'ensemble de la machine actuellement disponible. Cette demande tient compte de la répartition de nos calculs sur les SX8 de l'IDRIS et du CCRT. Afin de mener le prochain exercice GIEC en tant que groupe leader au niveau international nous avons besoin de l'assurance de bénéficier de cet accroissemen**t dès mi-2008.**

Un rapide calcul permet d'estimer le nombre de processeurs scalaires « équivalents » à nos besoins calcul soit 800 à 1000 processeurs de type Itanium ou 1300 à 1400 processeurs de type Opteron. Sur ces machines, de nombreuses incertitudes sur les outils subsistent : environnement, les performances, les temps de restitution des expériences complètes.

De plus, le plan de route montre l'ampleur des travaux scientifiques en cours qui visent tous à améliorer la climatologie des modèles et la crédibilité scientifique de l'IPSL. Mettre en priorité le portage sur un nouveau type de machine entraînera de facto un retard sur les développements que nous pourrons mener. Basculer la réalisation des simulations GIEC pour l'AR5 du NEC SX8 sur une nouvelle machine massivement parallèle nous paraît prématuré et dangereux.

Nous avons besoin de continuité dans les accès aux ressources calculs sur les 4 prochaines années.

II. Stratégie de parallélisation de codes climat de l'IPSL

Les performances des codes climatiques de l'IPSL sur les architectures informatiques sont au centre de nos préoccupations. Une stratégie de parallélisation de toutes les composantes du modèle système climatique a débuté il y a 13 ans. Des efforts conséquents ont porté sur l'implémentation de plusieurs techniques de supercalculs, destinées à exploiter au maximum la puissance de calcul mise à notre disposition, afin d'accélérer les temps de restitution de nos codes.

La parallélisation du modèle d'océan OPA a débuté dés 1988, mais avec des méthodes spécifiques aux calculateurs Cray. Une parallélisation plus générique a débuté en 1993 grâce à des collaborations avec l'ONERA, et l'utilisation scientifique n'a pas cessé depuis 1997. Le modèle de glace de mer LIM inclut dans OPA a été parallélisé en 2004 et 2005. En 2000, nous avons commencé les études sur la parallélisation du modèle d'atmosphère LMDZ, avec des collaborations avec les équipes des centres de calcul. Cette activité a été renforcée en 2005. La version parallèle de LMDZ est en cours de déploiement sur les calculateurs de l'IDRIS et du CCRT. La parallélisation du modèle de surfaces continentales ORCHIDEE a débuté en 2005, et est en cours d'achèvement. Début 2006, des travaux ont permis l'utilisation du coupleur OASIS 3 sur quelques processeurs. Le CERFACS a entamé en 2002 la réécriture complète du coupleur OASIS pour développer la version OASIS 4 entièrement parallèle. Enfin, depuis 2006 le projet CICLE (ANR calcul intensif) qui réunit l'IPSL le CERFACS et Météo-France a permis d'affecter du personnel à la parallélisation du système complet. Une première version du modèle couplant océan, glace de mer, atmosphère et surface continentales avec toutes les composantes parallèles est en cours de test.

Nos codes ont des échelles de temps de développement de l'ordre de la décennie. Ils capitalisent une expertise technique et scientifique acquise par une communauté vaste de développeurs et d'utilisateurs. Nous avons donc choisi des techniques de supercalcul pérennes et fiables, correspondant à des standards qui ne soient pas liés à un constructeur particulier. Leur mise en œuvre est faite de façon à permettre au plus grand nombre possible d'utilisateurs de participer aux développements.

Nous avons par ailleurs plusieurs contraintes propres :

- Les échelles de temps que nous traitons sont très longues, et nous contraignent à limiter fortement la résolution spatiale des modèles. C'est un obstacle majeur à la parallélisation massive des codes.
- Nous travaillons sur la sphère, ce qui impose des maillages avec des points singuliers, et donc des traitements spécifiques. La propagation d'ondes rapides dans l'atmosphère et les contraintes de conservation rend difficile, et non résolue aujourd'hui, l'utilisation de maillages composites sans points singuliers.
- Le temps de calcul est très réparti dans les codes, ce qui alourdit toute optimisation et interdit de faire du code spécifique à un calculateur particulier.
- Le portage et l'optimisation du coupleur Oasis est à garantir sur chaque nouvelle machine.
- Des soucis pour traiter les entrées/sorties sont toujours prévisibles (optimisation, recombinaison des sous-fichiers produits en parallèle, création des séries temporelles, atlas, …) et plus complexes encore sur des machines massivement parallèles dont l'environnement matériel et logiciel est encore peu standardisé.

- Le temps de développement des modèles climat représente plusieurs centaines d'homme-ans et s'étale sur des périodes très longues. Les différentes versions du modèle climat IPSL ont été figées en 1995 pour IPSL-CM0, en 1997 pour IPSL-CM1 et en 2005 pour IPSL-CM4.
- Le temps de restitution d'une simulation contraint l'exploitation scientifique. Il n'est pas équivalent d'avoir 6 résultats de simulations tous les 6 mois, et d'en avoir 1 chaque mois, ce qui permet une analyse des résultats intermédiaires .

Afin d'assurer la pérennité de nos applications, nous devons maintenir la portabilité de nos outils en généralisant les solutions optimales. Pour cela nous avons besoin de personnel efficace, compétent et connaissant en détail nos applications et les calculateurs disponibles dans les centres de calcul.

II.1. Les 4 niveaux de parallélisation

Nous exploitons aujourd'hui plusieurs méthodes d'accélération liées au parallélisme :

- Vectorisation des boucles de calculs. La vectorisation utilise la technique de 'pipelining' qui accélère les calculs, et d'ailleurs implantée dans certains processeurs scalaires. C'est aussi un premier niveau de parallélisation, puisqu'un processeur vectoriel de SX8R exécute 16 opérations en parallèle sans pénaliser le réseau mémoire, et peut atteindre 32 opérations simultanées.
- Accélération par découpage du code en tâches exécutées de manière concourante, mise en place à l'aide de directives OPENMP. Cette technique permet d'utiliser efficacement les architectures ou plusieurs processeurs partagent la même mémoire.
- Accélération par décomposition en sous-domaines, les échanges d'informations aux interfaces étant effectués par appels à la bibliothèque MPI. On communique alors entre processeurs n'ayant pas de mémoire commune.
- Accélération en exécutant simultanément les différents codes du modèle couplé sur différents processeurs.

Ces différents niveaux sont utilisés dans nos codes. Ils sont opérationnels, sauf pour OpenMP qui est dans le modèle d'atmosphère à l'état de prototype.

Il existe d'autres techniques complémentaires (microtasking, langages parallèles, …). Nous ne les utilisons pas car elles sont très peu répandues et par conséquent pas portables.

II.1.1. Vectorisation

Dans les codes climat, on distingue globalement deux composantes relativement distinctes et indépendantes. D'une part, la modélisation des processus dynamiques qui consiste à résoudre les équations hydrodynamiques de circulation tri-dimensionnelle de l'atmosphère ou de l'océan, et d'autre part la modélisation des processus physiques qui se charge de calculer l'ensemble des phénomènes locaux (convection, radiation, nuages, échange au niveau des couches limites, processus de surface, etc.…). La simulation de l'ensemble de ces processus a un point commun : les mêmes opérations de calcul sont appliquées à l'identique sur chaque point de la grille qui couvre le globe terrestre. En particulier lors des processus physiques les calculs en chaque point géographique sont totalement indépendants les uns des autres : chaque opération de calcul est donc effectuée sur un vecteur de points.

Les codes climatiques ont donc un caractère intrinsèquement vectoriel qui leur confère un net avantage, en terme de performances, en étant exécutés sur des machines dites « vectorielles ». En effet, ces machines sont justement conçues pour effectuer de manière simultanée un même calcul sur un vecteur de données, d'où des performances hors du commun lorsque le code est bien vectorisé. Le secret des performances de ces machines consiste en une parfaite adéquation entre la puissance de calcul des unités CPU et l'accès à la mémoire.

Les machines à base de processeurs dit « scalaires » (Itanium, Opteron) sont également capables de bonnes performances dans la mesure où la mémoire est peu sollicitée et le code tient en mémoire cache, qui est une mémoire très rapide mais de taille réduite. Ceci réduit les accès à la mémoire principale, plus lente. Dans le cas de nos codes, le rapport des performances se situe entre un facteur 5 (Itanium 16 Mo de cache) à 9 (Opteron 4 Mo de cache) en faveur d'un processeur vectoriel (NEC SX8R) que confirment des résultats observés sur des simulations à faible résolution (grille 4°). Néanmoins lors du passage à des résolutions plus élevées (grille 1°), les données ne tiennent plus en cache et l'accès à la mémoire devient un goulet d'étranglement important pour les processeurs scalaires, et les performances se dégradent : facteur 8 (Itanium) à 23 (Opteron) en faveur d'un processeur vectoriel.

C'est la raison pour laquelle la plupart des grands centres dédiés au climat ou à la météo sont équipés de manière conséquente de machines à la fois vectorielles et parallèles : Earth Simulator (Japon, NEC SX6), Oak-Ridge (USA : Cray X1E), Met Office (UK, NEC SX8), DKRZ (Allemagne, NEC SX6), Météo-France (NEC SX8R), HLRS (Allemagne, NEC SX8). Le projet HEcTOR au Royaume-Uni vient lui de choisir un calculateur Cray hybride (mélangeant processeurs scalaires et vectoriels).

Historiquement, la vectorisation a été la première forme de parallélisme de données exploitée par les codes de l'IPSL. Comme leur structure algorithmique est proche des concepts vectoriels, la mise en œuvre de cette technologie est simple et très peu invasive, ce qui allège considérablement le travail de mise au point et laisse au code toute sa lisibilité. D'ailleurs nos codes ne contiennent eux-mêmes aucune directive de vectorisation. L'ensemble de nos codes affiche un taux de vectorisation aux alentours de 99%, nous permettant de profiter au maximum des performances des machines vectorielles mises à notre disposition au CCRT et à l'IDRIS.

Nous retravaillons régulièrement l'optimisation des performances vectorielles, en particulier à l'arrivée de nouvelles machines.

II.1.2. Parallélisation en mémoire distribuée

Actuellement toutes les machines de supercalcul sont massivement parallèles, c'est-à-dire constituées d'un ensemble de nœuds de calcul relié par un réseau rapide permettant les échanges de données dans des temps de l'ordre de la microseconde.

Pour tirer parti de ces architectures, nos codes sont conçus pour exploiter la technologie de parallélisme MPI (Message Passing Interface). Le principe consiste à découper la grille de calcul en sous-domaines disjoints, qui sont assignés aux différents processeurs de calcul de la machine. Les données aux frontières des domaines, qui sont nécessaires à la poursuite du calcul, sont transférées d'un processeur à l'autre à l'aide de communications MPI.

Théoriquement, cette technologie permet d'exécuter des codes sur plusieurs milliers de processeurs. Il est toutefois nécessaire d'avoir des résolutions de grille de calcul extrêmement fines, de façon à éviter que le poids des communications MPI ne soit supérieur au temps de calcul des domaines. Il est également préférable d'avoir un algorithme qui permette un découpage tridimensionnel du domaine de façon à minimiser la taille des frontières entre sous-

domaines (minimisation du coût des transferts MPI), ainsi qu'une parfaite répartition de la charge de calcul (le calcul de chaque sous-domaine prend le même temps).

De par leur nature, les codes de l'IPSL ne répondent pas parfaitement à ces critères. En général, nos simulations couvrent des périodes de plusieurs siècles, avec des pas de temps (delta-t) de l'ordre de la minute. L'ensemble des processus calculés à chaque pas étant très consommateur en calcul, nous en sommes donc réduits à utiliser des grilles de faibles résolutions, afin que les temps de restitution des simulations restent raisonnables (2~3 mois). Nos grilles actuelles sont de l'ordre de 4°, à savoir 96x72x19 ce qui représente environ 100 000 mailles. A de si faibles résolutions maintenir une bonne 'scalabilité' au-delà de quelques dizaines de processeurs est une vraie gageure.

Par ailleurs, la résolution de nos modèles climatiques sur la sphère nous impose des contraintes algorithmiques fortes. Un certain nombre de quantités numériques comme l'énergie, le moment cinétique, l'enstrophie, etc… a besoin d'être conservé au cours du temps, sur des siècles, voire sur des millénaires (dans le cas du paléoclimat). Ces conservations sont indispensables pour simuler un bon équilibre climatique. Nous utilisons un schéma numérique à différences finies centré en longitude-latitude simple et robuste qui conserve l'ensemble de ces quantités.

La projection de la sphère sur une telle grille produit des points singuliers aux pôles qui conduisent à un resserrement des mailles, sur lesquelles les conditions CFL ne sont plus satisfaites. Pour les rétablir, on applique alors un filtre (de type FFT) sur les régions polaires (1/3 de la surface du globe). La parallélisation de ce type d'algorithme est délicate et nous impose un découpage en domaine en bande de latitude (voir Figure 3).

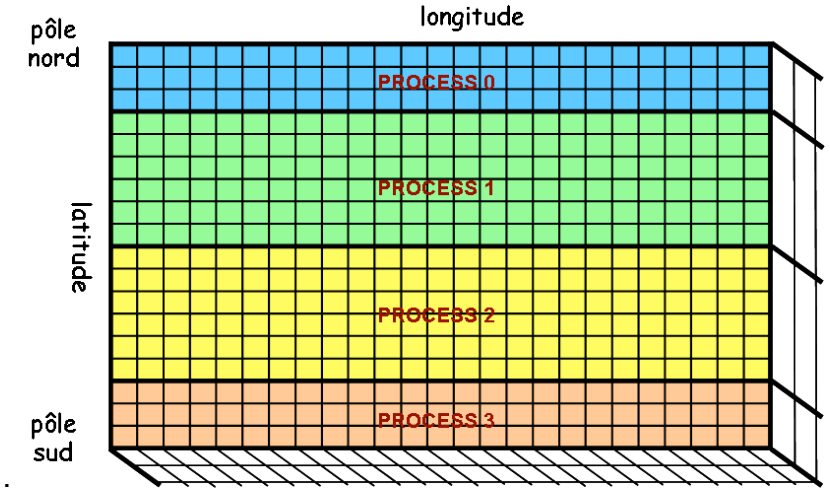

Figure 3 : découpage du domaine de LMDZ pour une parallélisation en mémoire distribuée

Pour permettre les échanges MPI, il est nécessaire de prévoir des halos de recouvrement (Figure 4), ce qui implique pour nos modèles que la taille des sous-domaines doit être au minimum de 3 bandes de latitude. Pour les résolutions actuelles (72 bandes de latitudes), le nombre maximal de processeurs sur le calcul est ainsi limité à 24. Mais compte tenu du déséquilibre de charge produit par le filtre sur les régions polaires, 16 processeurs est la limite pratique à cette résolution. Bien entendu, une résolution plus fine nous permet d'utiliser un plus grand nombre de processeurs. Pour le NEC SX-8 le nombre de processeurs recommandé est de 8 pour la première résolution et de 16 pour la deuxième pour rester au delà de 60%

d'efficacité du parallélisme (voir la section II.4). Il reste aussi du travail à faire sur la vectorisation. Les courbes de performance relatives pour la résolution 1°x1° sur 3 types de processeurs sont visibles dans la Figure 6. A cette résolution 1 processeur SX8R équivaut à 10 processeurs Itanium, et 4 processeurs SX8R à 60 Itanium.

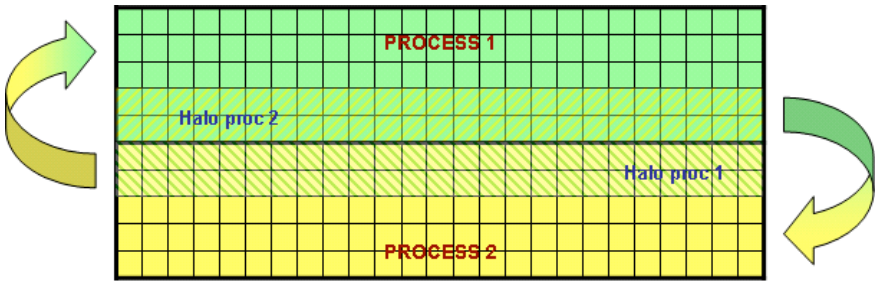

Figure 4 : principe des halos de recouvrement dans LMDZ.

II.1.3. Parallélisation en mémoire partagée

L'ajout d'un niveau de parallélisme supplémentaire nous permet de tirer parti des architectures SMP (plusieurs processeurs par nœud de calcul accédant à une mémoire commune), qui concerne la majorité des machines actuelles (hormis Cray/XT3 et IBM/BlueGene). Par ailleurs, l'arrivée massive des processeurs multi-cœurs, particulièrement bien adaptés à cette technologie, nous montre que cet investissement sera payant.

La parallélisation en mémoire partagée est basée sur le multithreading, qu'utilise le standard OpenMP. Cela consiste à partager les boucles de calcul sur plusieurs tâches (thread) à l'aide de directives spécifiques. Le concept se rapproche de la vectorisation, ce qui fait que pour nos codes, l'implémentation du parallélisme OPENMP s'est faite assez simplement.

OpenMP est opérationnel dans NEMO/OPA. Il est à l'état de prototype dans LMDZ, et l'optimisation est en cours.

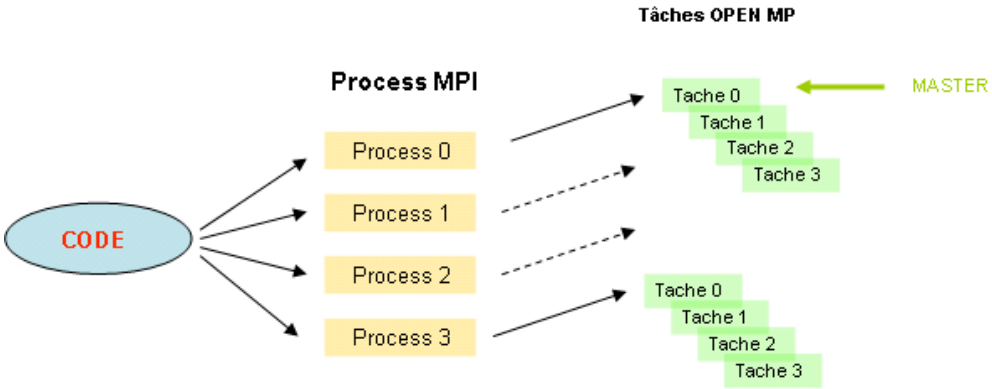

Figure 5 : emboîtement des parallélismes MPI et OpenMP.

II.1.4. Couplage

Le couplage ajoute un 4ème niveau de parallélisme au niveau des exécutables qui sont lancés simultanément sur des processeurs différents. Actuellement, nous couplons à l'aide de la technologie MPI-2, les codes LMDZ/ORCHIDEE avec NEMO/OPA et NEMO/LIM. La communication et les calculs auxiliaires (projections, conventions géographiques) sont alors effectués par le coupleur (OASIS) qui est un exécutable autonome. Pour illustrer les diffi-

cultés techniques du couplage, rappelons que MPI-2 n'est pas disponible sur les machines IBM. Le couplage se fait alors par MPI-1, ce qui rajoute des difficultés supplémentaires.

II.2. Etat des codes développés à l'IPSL

Tous nos codes bénéficient des différents niveaux de parallélisme. Il reste pour certains codes du travail pour passer du prototype à l'opérationnel.

II.2.1. NEMO

NEMO est un modèle de circulation générale de l'océan en point de grille (NEMO/OPA) comprenant un module de la glace de mer (NEMO/LIM) et un module de biogéochimie marine (NEMO/PISCES). Son maillage est déformé pour éviter les singularités des pôles. Pour le pôle Nord, la grille s'appuie sur des ellipses dont les deux points singuliers sont sur les continents, naturellement ou artificiellement. Le pôle Sud est au milieu du continent Antarctique. Ces points continentaux sont masqués dans les calculs. Il inclut les trois niveaux de parallélisation : vectorisation, OpenMP et MPI. Ce modèle est utilisé sur différents types de machines et est utilisé en opérationnel par Mercator-océan

II.2.2. LMDZ et ORCHIDEE

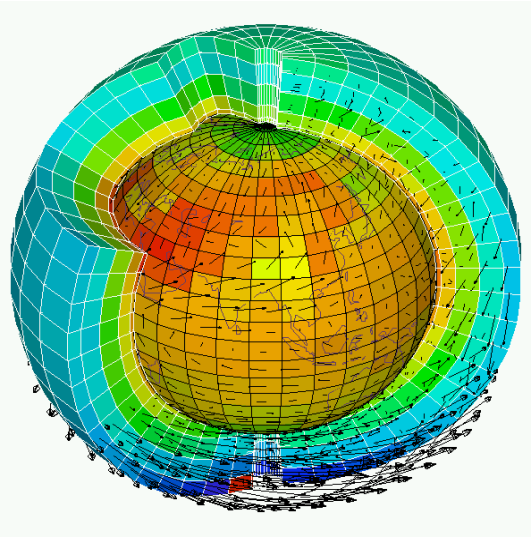

Figure 6 : maillage de la sphère pour LMDZ

LMDZ est un modèle atmosphérique à point de grille rectangulaire projeté sur le globe. Sa parallélisation se fait de façon différente selon les parties du modèle. La partie dynamique nécessite de connaître les informations globalement pour les opérations de filtrage aux pôles, pour les opérations de transport et pour l'utilisation de certains schémas numériques en différences finies. Pour optimiser l'équilibrage de charge, différents découpages de domaines sont utilisés dans un même pas de temps selon les parties de calcul. Les communications se font, dans cette partie en partageant les halos de recouvrement grâce aux appels à la bibliothèque MPI. De son côté, la physique est calculée sur des colonnes d'air et la parallélisation se fait alors en distribuant les points sur différents processeurs. LMDZ inclut informatiquement le modèle de surface continentale ORCHIDEE, dont la parallélisation suit le même découpage en domaine que la partie physique de LMDZ.

Les deux codes sont très bien vectorisés et sont totalement parallélisés avec MPI. Il existe à l'heure actuelle une version mixte MPI/OPENMP de LMDZ qui fonctionne avec succès dans un prototype. Un travail supplémentaire est nécessaire pour rendre ce niveau de parallélisation opérationnel.

II.2.3. INCA

INCA est le modèle de chimie atmosphérique et d'aérosols. Il s'agit en fait du couplage entre le modèle d'atmosphère LMDZ, un module de transport des espèces chimiques et un ensemble de réactions chimiques. Sa parallélisation reprend les paradigmes de la parallélisation de LMDZ. La parallélisation MPI est opérationnelle. La parallélisation OpenMP est à faire.

II.2.4. Couplé océan-atmosphère

Le modèle couplé s'appuie sur l'exécution simultanée de trois exécutables, chacun pouvant être exécuté sur plusieurs processeurs. Cette fonctionnalité n'est pas encore banale dans les environnements des calculateurs scientifiques et l'expérience a montré qu'il était toujours utile de prévoir un test de ses fonctionnalités lors des changements de machines pour trouver les bons réglages du système d'exploitation (voir MPI-1 obligatoire sur IBM). Il est en effet nécessaire de vérifier que ce couplage par OASIS fonctionne, que les réglages des classes d'exécution des travaux permettent son utilisation et que son utilisation routinière par plusieurs utilisateurs simultanés est possible. Il nous est impossible de réaliser les études climatiques actuelles sans cette fonctionnalité de couplage.

Un prototype du modèle couplé avec toutes ses composantes parallèles est aujourd'hui en test à l'IDRIS et au CCRT. Le passage en production est prévu dans le courant 2007.

II.2.5. Gestion des simulations : recombinaisons, post-traitements, etc. …

Une difficulté majeure pour nos codes parallèles est la gestion des entrées-sorties. Il n'existe pas aujourd'hui de bibliothèque d'entrée-sorties standard portable, adaptée à nos besoins, et performante. Chaque sous domaine de calcul (correspondant à un processeur ou un groupe de processeurs) écrit ses données. Le nombre de fichiers créés est ainsi multiplié par le nombre de processeurs utilisés et leur manipulation en est alourdie d'autant.

Un utilitaire particulier de recombinaison (rebuild) permet de reconstruire le fichier pour le domaine complet qui sera utilisé pour les analyses et post-traitement. Cette phase de recombinaison génère un volume très important d'entrée-sorties et constitue un goulet d'étranglement. Cette étape peut être extrêmement pénalisante et entraîner un ralentissement important de toute la chaîne de calcul. Cette partie des calculs doit rester sur des machines autonomes, moins chères que les supercalculateurs, ayant des accès rapides aux systèmes de fichiers et faisant tourner rapidement les outils nécessaires (rebuild, nco).

La réalisation des simulations est suivie de post-traitements systématiques : recombinaison des fichiers intermédiaires, création des séries temporelles, réductions de données, atlas. La aussi, nous avons besoin de machines avec des accès rapides aux systèmes de stockages.

Notre objectif est d'obtenir dans les délais impartis les simulations voulues. L'unité de mesure à retenir est l'année de simulation réalisée, fichiers recombinés et analyses inclus. Nous nous rapprochons ainsi des contraintes opérationnelles des services météorologiques.

II.3. Les étapes nécessaires pour passer en production scientifique

Pour passer en mode de production scientifique avec le modèle couplé parallèle sur un calculateur vectoriel/parallèle SX8, il reste plusieurs points à finaliser.

- Il faut finaliser les versions LMDZ4v3 et ORCHIDEE 2. Le parallélisme est implanté dans ces versions, mais des ajustements scientifiques restent à faire ;
- Nous avons identifié un goulet d'étranglement lors de la recombinaison des fichiers intermédiaires produits par les calculs ;
- Il faut prévoir une formation des utilisateurs. Nous considérons qu'il faudra un an à un an et demi pour que tous les projets scientifiques puissent utiliser le parallélisme.

Tout ceci implique également les scientifiques car l'expertise scientifique est vitale lors de toute évolution majeure des codes pour qu'ils soient physiquement validés. C'est crucial pour

chaque composante du système et pour le modèle couplé lui-même qui met en œuvre des rétroactions complémentaires.

Ces étapes (prévues en 2007) permettront d'utiliser de manière routinière une dizaine de processeurs pour chaque simulation. Les contraintes liées à l'utilisation d'un nombre plus élevé de processeurs, en particulier la parallélisation du coupleur, sont abordées dans la partie suivante.

II.4. Performances actuelles

Cette section fait état des performances des différents modèles et du couplé. Les comportements des différents modèles sont similaires rapportet montrent des accélérations en fonction du nombre de processeurs utilisés, accélérations plafonnées à des valeurs différentes selon les architectures de machines (vitesse des accès mémoire, vitesse des réseaux de communications, …). Les tests sont la plupart du temps avec des Entrées/Sorties minimales et constituent des valeurs indicatives. Ils ne sont pas représentatifs des simulations longues de type GIEC auxquelles sont ajoutées de nombreuses sorties diagnostiques puisque ces simulations lourdes sont ensuite analysées par des scientifiques aux intérêts très divers.

Atmosphère (LMDZ)

Le code LMDZ parallèle a été testé sur les machines à notre disposition : NEC SX8R, Itanium II quadrisocket bicoeur et Opteron quadrisocket 2.4 GHz. Le Tableau 4 et le Tableau 5 comparent les performances sur ces machines à deux résolutions. On y indique pour 3 types de machines l'augmentation de vitesse avec le nombre de processeur.

La résolution 96x72x19 (3,75°x2,5°x19 niveaux) est celle utilisée par l'IPSL pour le rapport AR4 du GIEC. Cette résolution sera utilisée pour l'AR5 pour les simulations couplant climat et cycles. A cette résolution 1 processeur SX8R équivaut à 6 processeurs Itanium, et 4 processeurs SX8R à 20 Itanium.

La résolution 360x180x19 (1°x1°x19) est celle que nous utiliserons pour améliorer la représentation des passages dépressionnaires aux moyennes latitudes. A cette résolution 1 processeur SX8R équivaut à 10 processeurs Itanium, et 4 processeurs SX8R à 60 Itanium.

	SX8R		**Itanium II**		**Opteron**	
	Temps (jours)	Accélération	Temps (jours)	Accélération	Temps (jours)	Accélération
1 proc.	13.8 j	1	55 j	1	116 j	1
2 proc.	7.2 j	1.9	32.5 j	1.7		
4 proc.	4.1 j	3.4	20.3 j	2.7	37.5 j	3.1
8 proc.	**2.5 j**	**5.4**	10.3 j	5.3	18.6 j	6.2
16 proc.	1.5 j	9.1	5.91 j	9.3	11.32	10.6
32 proc.	1.1 j	12.8				

Tableau 4 : augmentation de vitesse avec le nombre de processeur. Résolution 96x72x19 (3,75°x2,5°x19 niveaux), temps donnés pour 100 ans de simulation. On voit qu'à cette résolution 1 processeur SX8R équivaut à 6 processeurs Itanium, et 4 processeurs SX8R à 20 Itanium.

La comparaison des chiffres montre que sur le NEC SX-8 le nombre recommandé de processeurs est de 8 pour la première résolution et de 16 pour la deuxième pour rester au delà de 60% d'efficacité du parallélisme. Les courbes de performance relatives pour la résolution 1°x1° sont illustrés dans la Figure 7. Il faut noter que pour les processeurs Itanium, nous n'avons utilisé que la moitié des processeurs de chaque nœud pour ne pas surcharger le réseau mémoire.En terme de scalabilité, l'Itanium affiche de meilleures performances que le SX8 ce qui fait que l'on obtient une accélération de 39 sur 64 processeurs. Cela s'explique bien par les effets de mémoire cache (16Mo/cœur) qui sont très importants sur l'Itanium. Sur un

processeur, le code tient très peu en cache d'où des écarts importants de performance avec le SX8 (facteur 10). En augmentant le nombre de processeurs, donc en diminuant la taille de chaque sous-domaine, les données tiennent de plus en plus facilement en cache et accède de moins en moins à la mémoire lente, d'où une accélération intrinsèque des performances, donnant lieu parfois à des facteurs superlinéaires d'accélération (ex. facteur 8.6 à 8 proc.). En revanche, sur les processeurs vectoriels, la diminution de la taille des sous-domaines réduit l'efficacité de la vectorisation, ce qui fait chuter les performances, d'où une scalabilité réduite au-delà de 32 processus. Un travail d'optimisation est actuellements en cours afin de rétablir une bonne scalabilité sur un grand nombre de processus dans le cas des processeurs vectoriels.

	SX8 R		**Itanium II**		**Opteron**	
	Temps (jours)	Accélération	Temps (jours)	Accélération	Temps (jours)	Accélération
1 proc.	29.6 j	1	283 j	1		
2 proc.	15.1 j	1,96	144 j	1,97		
4 proc.	8.2 j	3,6	83 j	3,41		
8 proc.	4.7 j	6,3	33 j	8,58		
16 proc.	**2.7 j**	**10,8**	18,25	15,3	50 j	
32 proc.	1.8 j	16,7	12,2 j	23,2	33.5 j	
64 proc.			7,26 j	39		

Tableau 5 : augmentation de vitesse avec le nombre de processeur. Résolution 360x180x19 (1°x1°x19 niveaux), temps donnés pour 10 ans de simulation. On voit qu'à cette résolution 1 processeur SX8R équivaut à 10 processeurs Itanium, et 4 processeurs SX8R à 60 Itanium.

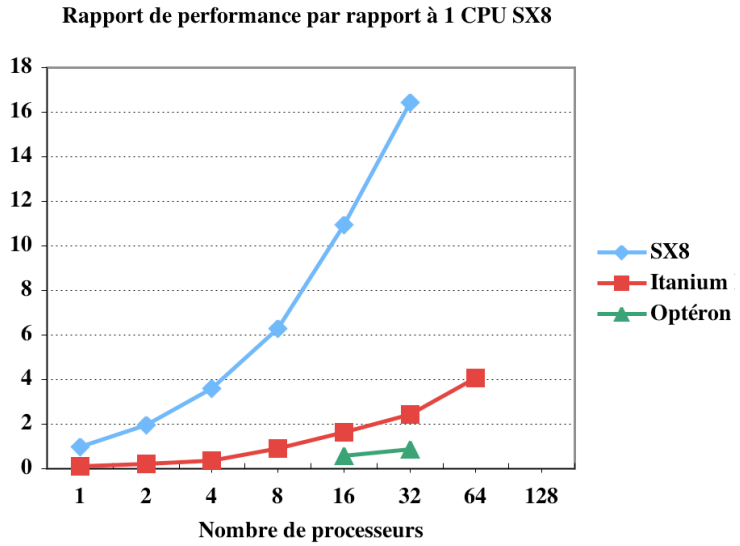

Figure 7 : courbes de performance en fonction du nombre de processeurs pour LMDZ en résolution 360x180x19. L'unité d'accélération est le rapport entre le temps de restitution du calcul sur un calculateur et le temps sur 1 processeur de SX8.

Océan (NEMO)

Le modèle d'océan NEMO est utilisé dans de nombreuses configurations sur différents types de calculateurs. L'expérience acquise permet d'évaluer les performances y compris à haute résolution. La configuration ORCA05 (0,5°x0,5°x31 niveaux ou 722x511x31) est la pro-

chaine configuration prévue pour le modèle climat de l'IPSL dès que les ressources calcul permettront de consacrer plus de ressources au modèle d'océan dans les simulations couplées.

Eléments de chiffrage en provenance de Drakkar, un projet utilisant NEMO sur Brodie (NEC SX-8) et Zahir (IBM)

Drakkar est un projet d'océanographie utilisant NEMO à différentes très hautes résolutions : 1615*1576*64, 1/12° sur l'Atlantique (NATL12) et 1442x1021x45, ¼° sur l'ensemble des océans (ORCA025). Les performances ont été mesurées sur Zahir (IBM SP de l'IDRIS, scalaire parallèle) et Uqbar (Nec SX8 de l'IDRIS, vectoriel parallèle).

En résumé, le tableau 6 montre que la configuration NATL12 est 5 fois plus lourde que la configuration ORCA025 en faisant le produit du nombre de points par le nombre de pas de temps calcul pour un an de simulation. Pour calculer un an de simulation, en prenant 6 fois plus de processeurs IBM dont l'efficacité par point de grille et par pas de temps est 16 fois moindre, le temps d'exécution diffère de 46%.

Le bilan montre ainsi que 186 processeurs IBM actuels sont équivalents à 10 processeurs NEC SX-8 pour cette application dans ce type de configuration.

Le détail des différentes configurations Drakkar : http://www.ifremer.fr/lpo/drakkar/embedded.htm

Configuration	NATL12	ORCA025	Rapport ORCA025/NATL12
Maillage	1615x1576x64	1442x1021x45	
Points calculés	100%	86% (domaines terre exclus)	
Nb points	$1.63.10^{8}$	$5.83.10^{7}$	0.36
Pas de temps	800 s	1440 s	1.8
Nb pas de temps pour un an de simulation	39420	21900	0.56
Mémoire	105 Go	94 Go	0.9
Ratio calcul	$=39420*1.6310^{8}$	$=21900*5.8310^{7}$	0.2
Machine	NEC SX-8	IBM Regatta	Rapport IBM/NEC
Temps CPU par point et par pas de temps	$3.87\ 10^{-7}$ s	$6.20\ 10^{-6}$ s	16.03
Temps CPU/an	690 h	2 200 h	5
Nb processeurs	32 procs (4 nœuds)	186 procs	5.81
Temps elapsed	20h	13.65 h	0.68
Efficacité	108%	87%	0.8

Tableau 6 : Eléments de chiffrage pour Drakkar.

Eléments de chiffrage de NEMO en configuration ORCA05L300 sur le Earth Simulator

Le modèle d'océan a aussi été porté sur le Earth Simulator au Japon (calculateur vectoriel et massivement parallèle de Yokohama). Il est utilisé seul ou couplé avec le modèle d'atmosphère ECHAM du Max-Planck Institut für Meterologie de Hambourg. Le Tableau 7 montre les performances atteintes par ORCA à très haute résolution (0,5°x0,5°x301 niveaux ou 722x511x301) pour différents nombres de processeurs sur le Earth Simulator. On observe un gain en vitesse d'un facteur 2,55 pour une multiplication par 4 du nombre de processeurs. Cette configuration est utilisée actuellement en routine avec 255 processeurs sur cette machine dans le cadre des coopérations CNRS-IFREMER-ESC et du projet INLOES ANR.

Nbre CPUS	64	73	85	102	128	170	255
Temps pour une année	6,04h	5,37h	4,77h	4,19h	3,60h	2,98h	2,37h
Années par jour	4	4,5	5	5,7	6,7	8	10

Tableau 7 : nombre d'années simulées par jour sur le Earth Simulator pour NEMO/OPA en configuration globale 0,5°x0,5°x300 niveaux.

La courbe d'accélération correspondante est montrée dans l'annexe V.3.

Modèle couplé océan-atmosphère IPSLCM4

Le modèle couplé océan atmosphère IPSLCM4 est en phase de tests de ses configurations parallèles. Les résultats sont disponibles uniquement sur NEC SX 8 (Tableau 8). Le portage sur d'autres machines est en cours. Là aussi, le portage du coupleur et son optimisation en constituent des étapes cruciales. Actuellement, l'océan prend 6% du temps calcul sur un processeur et 18% sur 4 processeurs. Ceci illustre la difficulté d'équilibrage de charge du modèle couplé nécessitant d'optimiser le nombre de processeurs pour chacune des composantes, ceci étant à faire pour chaque machine.

	1 proc	**4 proc**	
Temps écoulé	2 238 s	739 s	Soit un facteur 3 d'accélération
LMDZ 96x71x19	2 174 s	687 s 656 s 669 s 677 s = 2 689 s	Soit 23% de temps CPU supplémentaire
NEMO/OPA 182x149x31	130 s	129 s	
OASIS	40 s	43 s	
Total temps CPU	2 344 s	2 862 s	

Tableau 8 : Performance du modèle couplé IPSLCM4 en parallèle sur le SX8 R du CCRT, pour une simulation de 1 mois. Les temps différents sur chaque processeur montrent la difficulté de l'équilibrage, ce qui limite la scalabilité.

Une première expérience d'équilibrage du couplé avec toutes ses composantes parallèles a été réalisée sur le Earth Simulator en janvier 2006 par Arnaud Caubel. Le coupleur OASIS 3 était parallélisé en lançant plusieurs instances du coupleur en parallèle, chacune travaillant sur des champs différents, pour optimiser le couplé SINTEX (voir l'annexe V.4). Cette première expérience d'équilibrage sur 50 à 100 processeurs sera précieuse pour la mise en production parallèle du couplé IPSLCM4.

II.4.1. Comparaisons avec les groupes équivalents à l'étranger

En 1992, plusieurs organismes étatiques américains, dont le DOE (Department Of Energy) ont pris la décision d'abandonner les calculateurs vectoriels. Le DOE a alors lancé le programme CHAMMP, devenu plus tard Climate Change Prediction Program. Ce programme avait l'ambition de révolutionner la modélisation de la simulation climatique, avec de nouveaux algorithmes, de nouvelles grilles, etc. … Aucune des ambitions originales n'a été tenue. Des nouvelles grilles ont été développées, mais il reste encore aujourd'hui des obstacles numériques à leur utilisation dans des modèles de climat.

Un travail de parallélisation a été mené sur les différentes composantes du modèle de climat, en conservant pour l'essentiel les algorithmes et les grilles conçus dans les années 1970. L'investissement humain sur ces codes a été très important. Nous montrons ici les résultats de deux modèles résultant de ce travail.

FOAM

FOAM est un modèle américain de climat couplant un modèle atmosphérique de résolution 40 lats x 48 lons x 18 niveau (~9°x3,75°) et un modèle d'océan 128x128x24 (2,8°x1,4). Ceci est proche des résolutions que nous employons dans les études des climats du passé. FOAM utilise cependant une physique plus simple. Les méthodes de parallélisation sont très proches des notres.

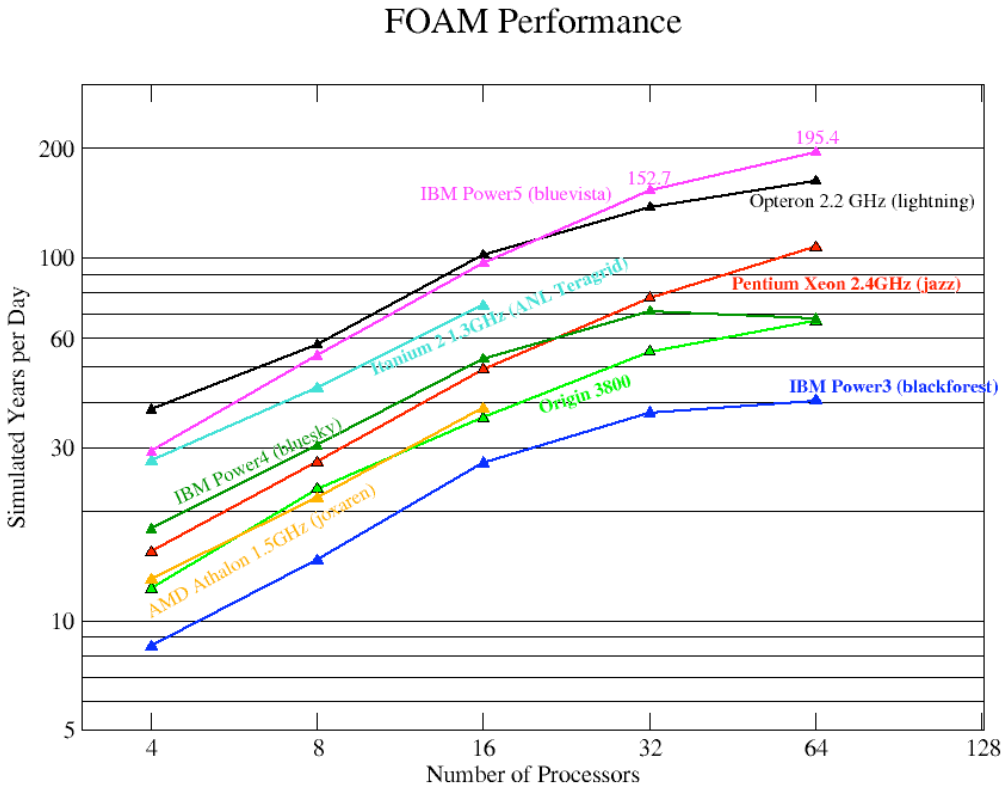

Figure 8 : mesures de performances de FOAM sur différents calculateurs. En abscisse, le nombre de processeurs employés, en ordonnée, le nombre d'années simulées par jour. dans le meilleur des cas, la multiplication par 16 du nombre de processeur ne conduit qu'à une accélération d'un facteur 5,6 (courbes verte et courbe rose).

Les courbes de la Figure 8 montrent pour différents calculateurs le nombre d'années simulées par jour en fonction du type de calculateur et du nombre de processeurs utilisés. Quel que soit le calculateur utilisé, la vitesse croit avec le nombre de processeur, mais les performances plafonnent lorsque l'on atteint 16 ou 32 processeurs.

Cette figure est une illustration de limites de scalabilité des modèles de climat. En effet, dans le meilleur des cas, la multiplication par 16 du nombre de processeur ne conduit qu'à une accélération d'un facteur 5,6 (courbes vert clair et courbe rose).

CCM3

CCM3 est le modèle de climat développé par le NCAR. La Figure 9 montre les performances en résolution T85 (~1,4°x1,4°), que le NCAR a utilisée pour les simulations du rapport AR4. L'IPSL utilisera une résolution de cet ordre pour le prochain exercice GIEC (AR5). En vert (pour un calculateur vectoriel Cray X1) et en rose (pour un calculateur scalaire IBM) les courbes idéales correspondant à une vitesse proportionnelle au nombre de processeurs. En rouge (pour le Cray X1) et en bleu (pour l'IBM), les performances effectivement atteintes. Les résultats obtenus sur ces codes par les équipes américaines après plusieurs années de tra-

vail montrent les limites intrinsèques de scalabilité des codes atmosphériques puisque les courbes mesurées (rouge et bleue) s'écartent des courbes idéales (verte et rose) au fur et à mesure que le nombre de processeurs utilisés croit. CCM3 a étcalculé utilisé par le NCAR pour ses simulations GIEC, à la fois sur le calculateur IBM de Oakridge (ORNL), ou les simulations restaient 3 mois en machine, et sur le Earth Simulator où le même type de simulations était réalisé en 3 semaines. Voir détails dans l'annexe V.2.

Il est également très instructif de consulter les plans du NCAR pour le prochain AR5 : http://www.ccsm.ucar.edu/about/ImpPlanCCSM4011207.doc

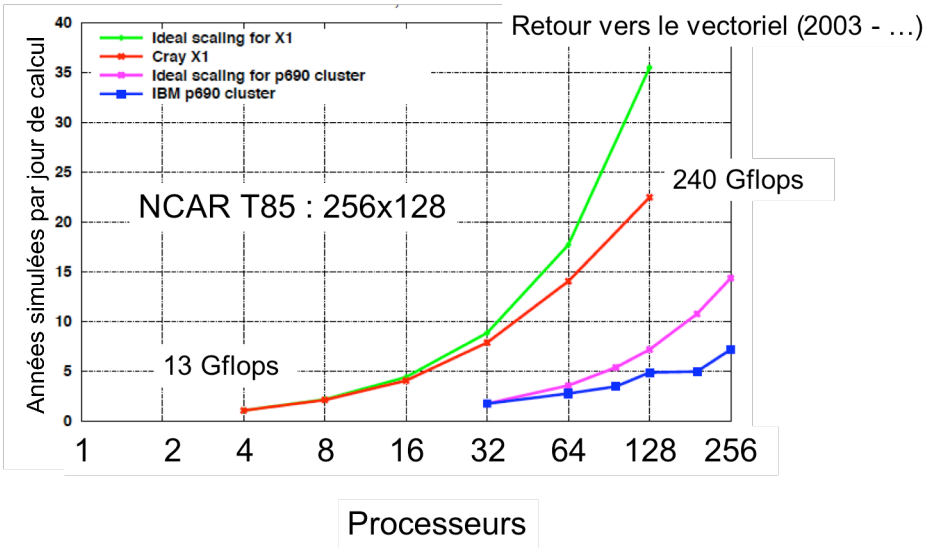

Figure 9 : performances du modèle CCSM3. En vert (pour un calculateur vectoriel Cray X1) et en rose (pour un calculateur scalaire IBM) les courbes idéales correspondant à vitesse proportionnelle au nombre de processeurs. En rouge (pour le Cray X1) et en bleu (pour l'IBM), les performances effectivement atteintes.

II.5. Conclusion

L'IPSL a développé ces dernières années une version parallèle de son modèle de climat et de ses composantes. La stratégie utilisée permettra d'utiliser un grand nombre de calculateurs d'architectures variées. Les méthodes utilisées et les performances atteintes sont équivalentes à ce qui se fait dans les autres groupes de modélisation climatique dans le monde. Le modèle actuel ainsi que son infrastructure sont parfaitement adaptés à l'utilisation optimale du SX8. Son utilisation opérationnelle commencera dans le courant de 2007.

Il reste plusieurs verrous à lever afin de pouvoir utiliser le modèle sur plusieurs dizaines, voire centaines de processeurs, et utiliser efficacement des calculateurs scalaires massivement parallèles.

III. Contraintes pour l'utilisation de machines massivement parallèles.

Nous avons expliqué dans la partie II.3 les étapes restant pour utiliser en routine une dizaine de processeurs par simulation. Nous étudions maintenant les étapes pour utiliser quelques dizaines de processeurs, puis quelques centaines. Par exemple, 255 processeurs du Earth Simulator (dans le cadre d'une convention internationale CNRS : MOU CNRS-IFREMER-ESC) sont utilisés par le modèle NEMO dans sa configuration ORCA05L300 (1/2°x1/2°x300 niveaux).

Dès à présent nous cherchons à assurer un portage de nos codes parallèles sur différents calculateurs représentatifs de différents types d'architectures parallèles. Nous nous préparons aussi à réaliser des simulations climatiques à très haute résolution sur un calculateur de classe exceptionnel qui pourrait voir le jour au niveau européen (projet HPC Europe). Les simulations climatiques impliquent de choisir le bon compromis entre l'augmentation de la résolution spatiale et l'augmentation de la durée des simulations et ce en fonction des ressources en calcul disponibles. Ainsi il est envisageable d'utiliser une très haute résolution pour une simulation météorologique à quelques jours, mais cette résolution devient prohibitive en temps de calcul si l'on doit simuler le climat pendant 300 ans. De plus, les grilles que nous utilisons et la représentation géographique intrinsèque de la Terre, qui oblige de filtrer l'écoulement aux pôles pour éviter les instabilités numériques, limitent les possibilités d'utiliser de façon optimale un grand nombre de processeurs.

Un travail supplémentaire est toujours nécessaire pour nous adapter à toute nouvelle machine. Deux points sont particulièrement critiques. D'une part, le portage et l'optimisation du modèle couplé, qui demande l'optimisation de 5 à 8 modèles suivant les configurations, est une étape lourde, car elle cumule l'ensemble des difficultés des différentes composantes auxquelles s'ajoutent les difficultés propres aux couplages. D'autre part, des soucis pour traiter les entrées/sorties sont toujours prévisibles (optimisation, recombinaison des sous-fichiers produits en parallèle, création des séries temporelles, atlas, …) et plus complexes encore sur des machines massivement parallèles dont l'environnement matériel et logiciel est encore peu standardisé. Afin d'assurer la pérennité de nos applications, nous devons maintenir la portabilité de nos outils en généralisant les solutions optimales. Quelque soit le calculateur parallèle envisagé, nos applications ne peuvent être efficaces que si la machine cible dispose de processeurs rapides, d'un réseau mémoire efficace, d'une mémoire suffisante par processeur, de connexions rapides …

III.1. Utilisation de quelques dizaines de processeurs

Les différentes composantes du modèle couplé (océan et atmosphère) sont capables d'utiliser chacune un nombre de processeurs qui selon la résolution employée va de la dizaine à la centaine. Dans de telles configurations, le coupleur représente un goulet d'étranglement important, car OASIS 3 n'est pas parallèle. Pour utiliser ces configurations sur le Earth Simulator, nous avons mis en place une pseudo parallélisation. Elle consiste à lancer plusieurs processus OASIS. Chacun des processus d'OASIS3 échange et interpole donc des champs de couplage différents. La situation optimale étant d'avoir autant de processus OASIS 3 que de champs de couplage à interpoler.

Cette « pseudo-parallélisation » d'OASIS 3 permet ainsi de réduire le temps de restitution des champs interpolés par OASIS 3 et atténue donc l'effet « goulet d'étranglement », en particulier lors de couplage de codes parallèles.

Cette technique est aujourd'hui exploitée en production pour les simulations de l'IPSL réalisées sur le Earth Simulator de Yokohama. Elle est aujourd'hui en cours de test sur les calculateurs de l'IDRIS et du CCRT. Ces tests permettront d'étudier l'équilibrage de charge. C'est en effet une phase délicate et spécifique à chaque calculateur et à chaque centre de calcul.

Recombinaison des fichiers

Lorsque l'on utilise plusieurs dizaines de processeurs, le travail de recombinaison des fichiers devient très lourd et pénalisant. A Yokohama, l'IPSL a installé une machine dédiée à la recombinaison des fichiers issus du Earth Simulator. Il reste un important travail à prévoir pour optimiser cette phase à l'IDRIS et au CCRT, qui pourrait demander l'installation de matériels spécifiques. Cette difficulté s'accroît avec le nombre de processeurs requis, puisqu'il faut recombiner tous les fichiers intermédiaires, ce qui constitue un goulet d'étranglement. Une autre solution consisterait en la manipulation en parallèle des fichiers intermédiaires par tous les utilitaires de post-traitement. C'est une autre piste à suivre.

III.2. Pour aller plus loin

Seul le modèle d'océan est capable aujourd'hui d'utiliser relativement efficacement une centaine ou plusieurs centaines de processeurs, ceci aux résolutions élevées utilisées aujourd'hui en océanographie. Ceci est dû aux expériences accumulées sur les T3E de l'IDRIS et du CEA dans les dernières années 90, puis sur le Earth Simulator depuis 2002. Sur cette dernière machine, les principales difficultés ont concerné l'environnement batch, les vitesses de transfert des fichiers entre le calculateur et les frontales, préliminaires à toute analyse, les recombinaisons de fichiers, les outils de débogage et d'analyse de performances.

Pour utiliser une centaine ou plusieurs centaines de processeurs avec le modèle d'atmosphère puis avec le modèle couplé océan-atmosphère, plusieurs étapes importantes restent à régler, qui sont détaillées ci-dessous : i) le parallélisme à mémoire partagé (OpenMP), ii) le passage à un nouveau coupleur, et iii) les entrées sorties.

i) OpenMP, un niveau de parallélisme supplémentaire

Une première piste explorée (voir II.1.3) consiste à ajouter un niveau de parallélisme dans LMDZ et ORCHIDEE. Cette piste est à creuser et à généraliser.

Nous avons choisi de paralléliser ainsi les boucles de calcul les plus externes, qui correspondent à la direction verticale. Cela permet, d'une part, d'être plus efficace en terme d'OPENMP et d'autre part de ne pas entrer en concurrence avec la vectorisation qui se fait elle sur les boucles de calcul les plus internes.

Nous disposons donc d'un modèle hybride MPI/OPENMP dans lequel chaque sous-domaine d'un processus MPI est lui-même découpé en plusieurs tâches. Sur la résolution standard (19 niveaux) on peut assigner 4 tâches supplémentaires sur les niveaux verticaux. Cela nous permet ainsi d'étendre la limitation du nombre de processeurs imposée par la parallélisation purement MPI.

La portée de la parallélisation OpenMP risque cependant d'être limitée par deux facteurs :

- Il s'agit d'un parallélisme en mémoire partagée. Beaucoup de calculateurs disposent d'un réseau mémoire ne permettant pas un débit soutenu lorsque plusieurs processeurs accèdent à la mémoire ;
- Peu de calculateurs bénéficient aujourd'hui d'une implémentation performante d'OpenMP ;

Nous aurons donc besoin d'une aide sur ce niveau de parallélisme, en particulier de la part des constructeurs de calculateurs.

ii) Un nouveau coupleur

Pour une utilisation du couplé sur machine massivement parallèle, nous devons passer au coupleur OASIS4, qui est conçu pour cet usage. Ceci est prévu dans le projet CICLE soutenu par l'ANR (appel d'offre GIGC05) et devrait être réalisé d'ici à la fin du projet fin 2008. Ce projet porte sur les développements d'OASIS 4 et sur son implémentation dans les codes de l'IPSL. Cet effort doit être maintenu dans le temps pour suivre l'évolution tant des modèles que des calculateurs.

iii) Les entrées/sorties

Le passage à plusieurs centaines de processeurs va mettre une pression très forte sur les systèmes d'entrée/sortie et de disques des calculateurs. On prévoit en particulier de gros soucis sur les machines massivement scalaires types clusters : débits en écriture, débits en lecture, masse de sous-fichiers intermédiaires, …

Nous avons besoin d'une aide sur ce point, ou beaucoup de solutions doivent être explorées. La solution adoptée sera probablement dépendante du calculateur. Un enjeu majeur sera de développer une solution pour encapsuler ces solutions spécifiques dans une interface simple et transparente pour des utilisateurs aux compétences informatiques hétérogènes.

III.3. Passer à des milliers de processeurs ?

L'utilisation d'un plus grand nombre de processeurs s'avère précieuse pour certaines études scientifiques (cas météorologiques, 'snapshots', océan à haute résolution), mais ne couvre pas et de loin l'ensemble des besoins de la communauté climat. Nous n'envisageons d'aborder la très haute résolution pour le climat qu'avec la perspective de pouvoir utiliser une machine européenne dimensionnée en conséquence. Dans ce cas, l'expérience a montré qu'il nous fallait accéder facilement à des ressources de taille similaire pour préparer les nouvelles configurations destinées à tourner in fine sur un calculateur exceptionnel. Ceci dimensionne l'écart entre les ressources facilement accessibles (nationales) et les ressources exceptionnelles (européennes) à un facteur 4 à 10 au maximum.

Enfin, la solution pour lever les limitations intrinsèques de nos applications est de travailler sur des nouvelles grilles pour la partie dynamique des codes pour éviter les points singuliers. Des grilles de type icosaédrique ou de type cube projeté sur la sphère conviendraient. L'expérience est actuellement en cours au Japon, en Allemagne et aux Etats-Unis, mais n'a pas encore d'application en climatologie, malgré des applications pour des simulations très courtes (échéances météorologiques de quelques jours). Cette nouvelle génération de modèles ne sera pas en production avant 10 à 15 ans. Nous commençons à mener une réflexion pour développer de tels modèles en parallèle de nos activités scientifiques actuelles. Il faut mettre en place un groupe de travail pour aborder les modèles de la prochaine génération, faisant appel à de nouvelles grilles, ce qui nécessite une approche multidisciplinaire alliant des compétences en mécanique de fluides géophysiques, algorithmique et numérique. Un renforcement à court terme des effectifs pour aborder pleinement ces nouveaux aspects est nécessaire. Nous devons préparer une stratégie de développement d'une telle grille ou de coopération avec d'autres groupes. C'est un travail de longue haleine qui passerait en production d'ici 10 à 15 ans et que nous devons commencer à mener en parallèle de nos activités scientifiques actuelles.

IV. Bibliographie

Ressources nécessaires pour la modélisation du climat à l'IPSL, Pôle de modélisation du climat de l'IPSL, Mai 2004 : http://dods.ipsl.jussieu.fr/mafoipsl/DOCUMENTS/NOTES/Pole_Prospective_calcul_2004.pdf

The new IPSL climate system model: IPSL-CM4, Marti O. et al., Note du pôle de modélisation IPSL N° 26 : ISSN 1288-1619. Voir : http://dods.ipsl.jussieu.fr/omamce/IPSLCM4/

Oasis : S. Valcke, 2006: OASIS3 User Guide (prism_2-5), PRISM Support Initiative Report No 3, CERFACS, Toulouse, France, 64 pp.

IPCC : The Intergovernmental Panel on Climate Change. Voir www.ipcc.ch. Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC).

Météo-France, Bi-ORAP 49, Météo France renforce ses moyens de calcul, Dominique Birman, http://www.irisa.fr/orap/Publications/Bi-orap/biorap-49.pdf

V. Annexes

V.1. Calendrier de réalisation des scénarios GIEC à l'IPSL pour AR4

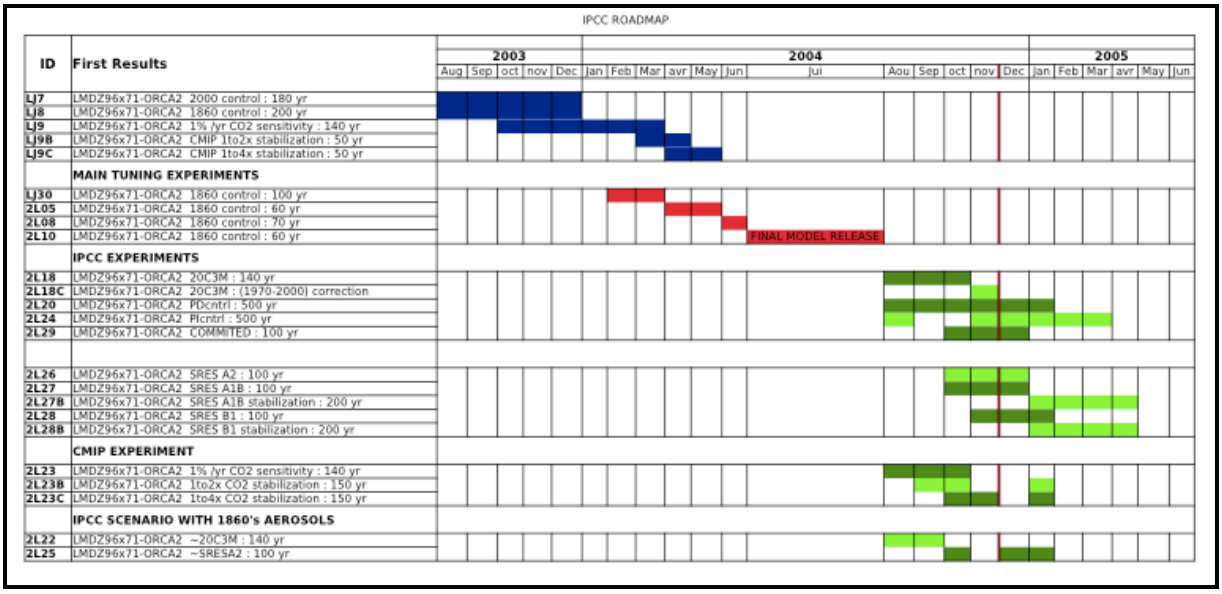

Tableau 9 : Calendrier de réalisation des scénarios GIEC à l'IPSL pour AR4.

Le Tableau 9 fait état du calendrier de réalisation des simulations de l'A4 à l'IPSL sur la période 2003-2005. Il ne prend pas en compte les simulations climat-carbone. Dans le même temps, à l'IDRIS, le calculateur NEC SX-5 Uqbar était toujours accessible puisqu'il a été installé en 2000 et qu'il a été arrêté à l'automne 2006. Le portage sur le NEC SX-8 Brodie s'est fait plutôt facilement durant l'été 2006, mais a tout de même nécessité 3 mois d'adaptation.

V.2. Calendrier de réalisation des scénarios GIEC du NCAR pour AR4

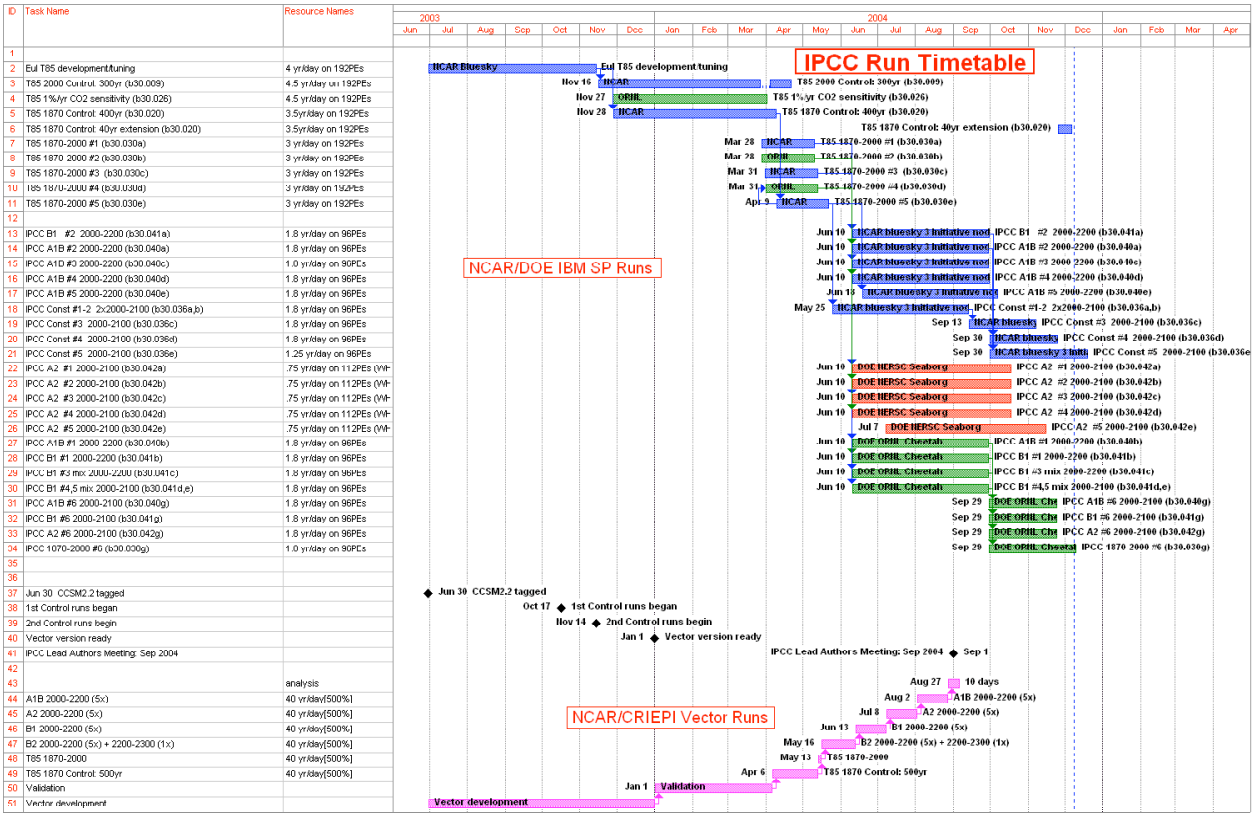

Tableau 10 : Calendrier de réalisation des scénarios GIEC du NCAR pour AR4.

Le Tableau 10 montre le calendrier des simulations AR4 réalisées par le NCAR . La démarche est similaire à celle de l'IPSL. Le NCAR disposant de moyens de calculs plus important que l'IPSL, il a réalisé plus de simulations, à des résolutions plus élevées. On constate que les simulations réalisées sur le Earth Simulator (en rose, ligne 44 par exemple) ont été réalisées en effectuant 40 ans de simulations par jour. Ceci est à comparer aux 9 années simulées par jour (5 fois 1.8 années par jour) sur les calculateurs scalaires du NCAR et de l'ORNL (en bleu, lignes 14 à 17).

V.3. Efficacité et accélération de la configuration ORCA05L300 (722x511x301) du modèle OPA sur le Earth Simulator

Le modèle d'océan ORCA05L300 (modèle global en résolution 722x511x301) est utilisé couramment sur le Earth Simulator. Une étude de scalabilité complète a été exigée par les equipes du Jamstec avant d'autoriser l'utilisation d'un grand nombre de processeurs. La résolution relativement élevée lui permet d'utiliser 260 processeurs de façon efficace. Passer de 64 à 255 processeurs permet une accélération d'un facteur 3,12.

OPA8.2 for ES: ORCA05 L 300 1728 time steps							
#CPU	PNxP	Tn	V.Op.R(%)	Total GFLOPS	Mem(GB)	hour/365d	GFLOPS/CPU
64	8x8	1431	99,61%	205	143	6,045	3,2137
73	3x8+7x7	1271	99,60%	232	149	5,369	3,1880
85	8x8+3x7	1130	99,58%	264	163	4,773	3,1170
102	11x8+2x7	991	99,56%	308	180	4,186	3,0208
128	16x8	852	99,55%	373	209	3,599	2,9145
170	16x8+6x7	706	99,50%	468	248	2,982	2,7544
255	31x1+7x1	562	99,45%	640	339	2,374	2,5124

Tableau 11 : mesure de performances parallèle pour le modèle d'océan

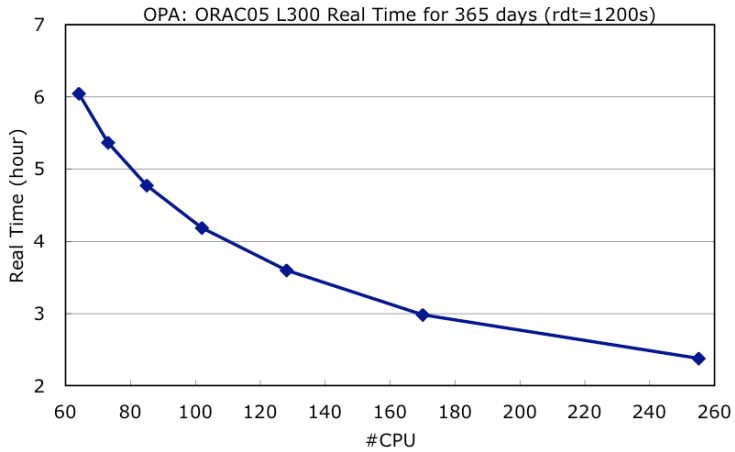

Figure 10 : courbe de scalabilité de ORCA05 sur Earth Simulator correspondant au Tableau 11.

V.4. Optimisation du nombre de processeurs utilisé sur le Earth Simulator par le couplé SINTEX

Le séjour d'Arnaud Caubel en 2006 au Japon lui a permis de mettre en œuvre la parallélisation du coupleur Oasis 3. Dans le cadre du projet européen SINTEX, le modèle utilisé couplait le modèle d'océan ORCA05 (722x511x300) et le modèle d'atmosphère ECHAM (résolution T106, soit 320x160). Oasis 3 a été parallélisé en utilisant plusieurs instances du coupleur. Voir le rapport complet :

http://dods.ipsl.jussieu.fr/projets_ES/INLOES/2006/CRM_2006_2.pdf

La Figure 11 résume les performances atteintes en fonction du nombre de processeurs utilisés et de la répartition des différents modèles sur ces processeurs. La configuration optimale est d'utiliser 4 processeurs pour Oasis, 8 processeurs pour Opa et 40 processeurs pour le modèle d'atmosphère ECHAM. Il est alors possible de simuler 3,5 années par jour en utilisant 52 processeurs du Earth Simulator qui en comporte 5120. Dans ce modèle, l'atmosphère est sur une grille T106L31 (maillage proche de 320x160x31) avec un pas de temps de 6 minutes. L'océan est sur une grille de 0.5°, 31 niveaux avec un pas de temps de 36 minutes.

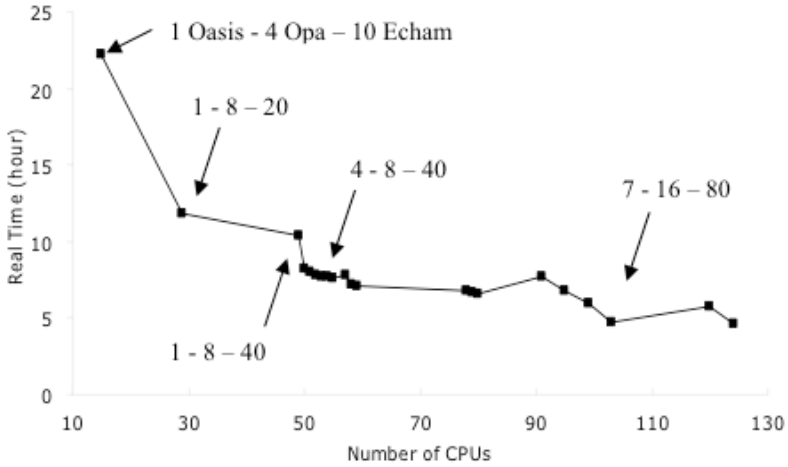

Figure 11 : tests d'optimisation du modèle couplé SINTEX sur Earth Simulator temps mesurés pour 1 an de simulation en fonction du nombre de processeurs et de leur répartition entre les modèles. La configuration à 52 processeurs est la plus efficace (4 pour Oasis, 8 pour OPA et 40 pour Echam).

VI. Table des matières

I. Plan de travail pour le GIEC 2009 2
I.1. Introduction 2
I.2. Plan de route pour le prochain exercice GIEC (2009) 2
I.2.1. Les enjeux scientifiques 2
I.2.2. Les versions des modèles 3
I.2.3. Le passage en parallèle pour tous les utilisateurs (de mi-2007 à mi-2008) 5
I.2.4. Besoins en calcul associés 7
I.3. Bilan des besoins pour l'AR5. 9
II. Stratégie de parallélisation de codes climat de l'IPSL 11
II.1. Les 4 niveaux de parallélisation 12
II.1.1. Vectorisation 12
II.1.2. Parallélisation en mémoire distribuée 13
II.1.3. Parallélisation en mémoire partagée 15
II.1.4. Couplage 15
II.2. Etat des codes développés à l'IPSL 16
II.2.1. NEMO 16
II.2.2. LMDZ et ORCHIDEE 16
II.2.3. INCA 16
II.2.4. Couplé océan-atmosphère 17
II.2.5. Gestion des simulations : recombinaisons, post-traitements, etc. … 17
II.3. Les étapes nécessaires pour passer en production scientifique 17
II.4. Performances actuelles 18
II.4.1. Comparaisons avec les groupes équivalents à l'étranger 21
II.5. Conclusion 23
III. Contraintes pour l'utilisation de machines massivement parallèles. 25
III.1. Utilisation de quelques dizaines de processeurs 25
III.2. Pour aller plus loin 26
III.3. Passer à des milliers de processeurs ? 27
IV. Bibliographie 28
V. Annexes 29
V.1. Calendrier de réalisation des scénarios GIEC à l'IPSL pour AR4 29
V.2. Calendrier de réalisation des scénarios GIEC du NCAR pour AR4 30
V.3. Efficacité et accélération de la configuration ORCA05L300 (722x511x301) du modèle OPA sur le Earth Simulator 30
V.4. Optimisation du nombre de processeurs utilisé sur le Earth Simulator par le couplé SINTEX 31
VI. Table des matières 33