

La politique française dans le domaine du calcul scientifique

*rapport à monsieur le ministre
de l'éducation nationale,
de l'enseignement supérieur
et de la recherche*

*à monsieur le ministre délégué
à la recherche*

Mars 2005

N° II-B-14-2004

N° 2005-017

La politique française
dans le domaine du calcul scientifique

Mars 2005

Emmanuel SARTORIUS
*Ingénieur général
des Télécommunications*

Michel HÉON
*Inspecteur général de l'administration
de l'éducation nationale et de la recherche*

SOMMAIRE

INTRODUCTION	1
I- POURQUOI LE CALCUL INTENSIF ?	2
II- LES GRANDS CENTRES DE CALCUL SCIENTIFIQUE EN FRANCE	3
LE CINES	3
L'IDRIS.....	4
LE CEA.....	4
LE CENTRE DE CALCUL DE L'IN2P3	5
SYNTHESE	6
III- LES MESOCENTRES ET LES MOYENS DES LABORATOIRES	7
IV- LE DIAGNOSTIC	8
V- LA FRANCE ET L'EUROPE	15
DEISA	15
LES GRILLES	16
LE 7 ^{EME} PCRD.....	16
L'INITIATIVE TRIPARTITE ALLEMAGNE - FRANCE - ROYAUME-UNI.....	16
VI- LES LOGICIELS	17
VII- BESOINS SCIENTIFIQUES ET BESOINS INDUSTRIELS	20
VIII- EXTERNALISER LE CALCUL SCIENTIFIQUE ?	21
IX- CALCUL VECTORIEL, CALCUL PARALLELE OU GRILLES ?	21
X- ELEMENTS D'UNE POLITIQUE	23
RECOMMANDATION 1 : METTRE EN PLACE UN COMITE STRATEGIQUE DU CALCUL SCIENTIFIQUE	24
RECOMMANDATION 2 : COMBLER LE RETARD FRANÇAIS EN CALCUL INTENSIF	25
RECOMMANDATION 3 : LA STRUCTURATION DES ACTEURS DU CALCUL INTENSIF	26
RECOMMANDATION 4 : RENFORCER LA COOPERATION EUROPEENNE	27
RECOMMANDATION 5 : FAIRE FACE A LA POSSIBLE DISPARITION DU CALCUL VECTORIEL	29
RECOMMANDATION 6 : DEVELOPPER LES SYNERGIES EN MATIERE DE LOGICIELS.....	29
RECOMMANDATION 7 : DEVELOPPER UN RESEAU D'EXPERTS	30
RECOMMANDATION 8 : ACCROITRE ET PERENNISER LES MOYENS FINANCIERS DU CALCUL INTENSIF	31
ANNEXES	33

Introduction

Par lettre du 21 septembre, le ministre délégué à la Recherche a demandé le concours de l'Inspection générale de l'administration de l'Education nationale et de la Recherche et du Conseil général des technologies de l'information pour animer un groupe de travail dont les travaux seraient destinés à définir et à préfigurer les structures de concertation et de décision nécessaires à la mise en œuvre d'une politique forte en matière de calcul scientifique en France (voir annexe 1).

Les rapporteurs ont donc tenu cinq réunions d'un groupe réunissant les représentants des administrations concernées, les directeurs des grands centres de calcul, des représentants des grands utilisateurs scientifiques et industriels (CNRS, CEA, EDF), ainsi que d'organismes liés étroitement au développement du calcul scientifique (INRIA et RENATER). L'annexe 2 donne la liste des membres du groupe de travail ainsi que celle des autres personnalités entendues par les rapporteurs.

Ceux-ci ont également tenu trois réunions ciblées sur la prospective avec des représentants du CNRS pour la première, de la communauté des nanotechnologies pour la seconde, et de la communauté de la biologie pour la troisième.

L'ensemble de ces réunions et de ces auditions ont permis de dégager un consensus sur la nécessité d'un pilotage stratégique du domaine, d'une meilleure coordination entre centres de calcul et d'un financement à un niveau convenable, et surtout régulier, du parc français de grands ordinateurs scientifiques. A partir de là les avis divergent sur les modalités de mise en œuvre de ces principes généraux. Il doit donc être clair que les opinions exprimées dans ce rapport et les recommandations qu'il contient sont celles des seuls rapporteurs.

Ceux-ci tiennent à remercier toutes les personnalités qu'ils ont rencontrées pour le temps qu'elles leur ont consacré et les échanges fructueux qu'elles ont eus avec eux. Ils remercient plus particulièrement les membres du groupe de travail qu'ils ont animé entre novembre 2004 et février 2005 qui, par leurs contributions, leurs remarques et leur connaissance approfondie du sujet, leur ont permis de progresser dans leurs travaux. Ils remercient enfin tout spécialement Alain Lichniewsky, chargé de mission à la direction de la Recherche, avec qui ils ont eu de nombreux échanges, qui les a fait bénéficier de son expérience et de sa connaissance des acteurs et qui a, en outre, assuré avec efficacité et dévouement le secrétariat du groupe de travail.

I- Pourquoi le calcul intensif ?

Désormais incontournables aussi bien dans le domaine de la recherche que dans celui de l'industrie, la modélisation et la simulation¹ sur ordinateur ont pour objet de créer une représentation virtuelle d'un objet, d'un phénomène ou d'un concept afin de l'analyser et de le comprendre, de façon à pouvoir prévoir son comportement, voire le contrôler. Modélisation et simulation se sont introduites naturellement entre théorie et expérience, qui sont au cœur de la démarche scientifique. Elles constituent désormais un moyen privilégié d'investigation pour les chercheurs. Elles sont devenues une condition nécessaire du progrès scientifique dans des domaines aussi variés que la climatologie, la chimie quantique, la biologie, la fusion contrôlée ou les nanotechnologies.

En outre, dans certains domaines, la nature même de la connaissance scientifique a évolué au cours des dernières décennies, en raison de l'étude de systèmes et de phénomènes complexes. Elle repose désormais sur la gestion et l'utilisation d'une masse d'information qu'il faut systématiser, recueillir, distribuer et exploiter. La biologie génomique et post-génomique, qui met en jeu des volumes d'informations toujours plus importants, constitue l'exemple le plus frappant de cette évolution.

Les enjeux économiques de la modélisation et de la numérisation sur ordinateur dépassent de beaucoup la mise à disposition des chercheurs d'outils particulièrement performants. D'une part la capacité d'un état à effectuer des modélisations et des simulations sur ordinateur est devenu un élément essentiel de sa crédibilité aussi bien dans le domaine de la défense (nucléaire) que dans celui de la préservation de l'environnement (climatologie). D'autre part, le niveau de compréhension et de caractérisation des systèmes complexes autorisé par la simulation numérique permet de réduire les risques, partant les coûts de développement, des grands programmes et d'optimiser les investissements industriels, dans les secteurs de pointe (microélectronique, nanotechnologies, sécurité des installations). La puissance de calcul dont dispose un pays est ainsi devenue un facteur essentiel à sa crédibilité sur la scène internationale, au succès de sa politique de recherche et un levier important pour sa compétitivité industrielle.

La mise en œuvre de ces techniques nécessite le recours à des calculateurs particulièrement puissants que leurs performances situent bien au-delà de celles des produits catalogue des constructeurs d'ordinateurs. C'est pourquoi on parle à leur propos de *calcul intensif* ou de *calcul de haute performance* (*High Performance Computing* ou *HPC*, en anglais).

La nécessité d'affiner sans cesse les modèles pour de meilleures prévisions (météorologie) ou d'étendre leur application dans le temps (océanographie) ou dans l'espace (nombre de molécules prises en compte dans une réaction chimique) conjuguée au progrès technologique conduit à améliorer sans cesse les performances des calculateurs. C'est ainsi que de nombreuses simulations numériques recourent à la technique du *maillage* qui consiste à découper l'objet étudié (l'atmosphère terrestre, un avion, etc.) en petits volumes élémentaires, les *mailles*. Augmenter la précision de la simulation signifie réduire la taille des mailles. Mais, si on divise celle-ci par 10, on multiplie le nombre de mailles par 1 000 (10^3). Or la puissance de calcul nécessaire évolue au moins dans les mêmes

¹ La simulation consiste à reproduire par le calcul, en général dans un souci d'économie ou parce que l'expérimentation est impossible, le fonctionnement d'un système dont on connaît les principes de fonctionnement (*crash tests* de voitures, simulateurs de vol, par ex.). En revanche, la modélisation consiste à essayer d'approcher de façon simplifiée, par choix délibéré ou par nécessité, par le calcul le fonctionnement de systèmes complexes (les océans, par ex.) ou d'évaluer les conséquences de théories scientifiques (la chromodynamique quantique, par ex.). La simulation numérique peut servir à valider ou à exploiter ces modèles. C'est souvent le transfert de codes de simulation (codes de simulation de *crashes* ou d'aérodynamique) qui concrétise le transfert des connaissances du monde de la recherche à celui de l'industrie.

proportions, si ce n'est plus vite, que le nombre de mailles, en dépit des progrès des sciences de l'algorithmique et du calcul.

Cette tendance conduit à l'utilisation d'ordinateurs de plus en plus puissants. Elle demande aussi un grand savoir-faire de la part des chercheurs qui doivent développer les programmes correspondants (les *codes*), suffisamment parallélisés pour exploiter au mieux ces machines. Celles-ci exigent une grande fiabilité pour pouvoir fonctionner sans faute pendant des centaines d'heures, voire des milliers dans certains cas extrêmes.

II- Les grands centres de calcul scientifique en France

La France a mis en place un dispositif national à trois niveaux pour donner aux chercheurs accès aux moyens de calcul, de stockage et de gestion de données dont ils ont besoin. Ce dispositif comprend :

- un premier niveau avec les quatre grands centres nationaux de calcul intensif que sont le CINES et l'IDRIS, le centre de calcul du CEA et celui de l'IN2P3² ;
- un échelon intermédiaire de *mésocentres*, généralistes ou thématiques, dans les régions ;
- les moyens propres des laboratoires et des universités.

Enfin, le réseau à haut débit RENATER assure l'interconnexion de l'ensemble, ainsi que l'accès des calculateurs français au réseau européen GEANT.

Le CINES

Créé en 1999, le CINES (Centre Informatique National de l'Enseignement Supérieur) est situé à Montpellier, où il a succédé au CNUSC (Centre National Universitaire Sud de Calcul). C'est un établissement public administratif (EPA) national placé sous la tutelle du ministère de l'Education nationale et de la recherche, avec une triple mission :

- de calcul numérique intensif ;
- d'exploitation de bases de données d'information et de documentation au profit des organismes de recherche publique ou des établissements d'enseignement supérieur qui le demandent ;
- d'expertise et formation en matière de réseaux informatiques ; le CINES collabore notamment avec le GIP RENATER dont il accueille une antenne.

² On n'a pas tenu compte ici des moyens de calcul de Météo France dans la mesure où ils sont dimensionnés en fonction des besoins opérationnels de l'établissement, même s'ils sont également utilisés pour des activités de recherche centrées sur l'amélioration des prévisions météorologiques.

Le CINES dispose principalement d'un ordinateur IBM SP 4 de 1,85 teraflop³ et d'un ordinateur SGI Origin 3 800 de 0,8 teraflop.

Le CINES emploie 50 personnes, dont 41 de l'Education nationale et 5 du CNRS.

L'IDRIS

L'IDRIS (Institut du Développement des Ressources en Informatique Scientifique) a été créé en 1993, à l'issue d'une profonde restructuration de l'informatique scientifique du CNRS, dont il est une unité propre de service (UPS). Administrativement, l'IDRIS est rattaché au département Sciences et Techniques de l'Information et de la Communication (STIC) du CNRS. Situé à Orsay, l'IDRIS intervient comme une structure de service qui assure la mise en place et l'exploitation d'un environnement de calcul intensif répondant au besoin de communautés scientifiques qui ont besoin de très grandes puissances de calcul. L'IDRIS dispose essentiellement d'un ordinateur vectoriel NEC SX-5 de 0,3 teraflop et d'un ordinateur parallèle IBM SP 4 de 6,5 teraflops.

Par ailleurs, l'IDRIS jouit également d'un certain *leadership* à l'échelle européenne en matière de réseaux de grands ordinateurs (voir p. 15).

L'IDRIS emploie une cinquantaine de personnes, essentiellement à statut CNRS.

Le CEA

Le Commissariat à l'Energie Atomique (CEA), pour sa part, a procédé il y a quelques années au regroupement de l'ensemble de ses moyens de calcul sur le seul centre de Bruyères-le-Châtel, qui relève de sa direction des applications militaires (DAM). La DAM exploite en fait deux ensembles de calcul bien distincts :

- des moyens militaires, particulièrement puissants mais radicalement isolés de l'extérieur (aucune connexion), qui concourent à la crédibilité de la dissuasion nucléaire française par les simulations qu'ils permettent après les arrêts des essais nucléaires dans le Pacifique ; ces moyens ne sont utilisés qu'à la marge pour certains besoins scientifiques très particuliers (génomique, par ex.) ; ils reposent actuellement sur un cluster HP SC 45 de 5,1 teraflops. A la suite d'un appel d'offres lancé en 2004, le CEA disposera dès fin 2005 d'une nouvelle machine de la classe des 60 teraflops.
- le Centre de Calcul de la Recherche Technologique (CCRT), qui regroupe les autres moyens de calcul du CEA ; le CCRT a pour vocation principale de satisfaire les besoins des pôles civils du CEA en matière de grands calculs scientifiques, mais aussi de pratiquer une

³ La puissance d'un ordinateur se mesure fréquemment en nombre d'opérations en virgule flottante par seconde (*floating point operation per second* en anglais) qu'il peut exécuter, d'où le nom de *flop* donné à cette unité. Ce sont surtout ses multiples qui sont utilisés : le gigaflop (10^9 flops), le teraflop (10^{12} flops) et le petaflop (10^{15} flops). Pour des applications n'effectuant pas ce type d'opérations, on recourt à d'autres unités de mesure qui reposent sur des benchmarks empiriques standardisés, tel le SPECint (cf. <http://www.spec.org/>). Sauf indication contraire, les puissances indiquées dans ce rapport sont des puissances *crête*, c'est-à-dire la puissance maximum que peut délivrer la machine. La puissance crête est évidemment supérieure, souvent d'un facteur 5 à 10, à la puissance *soutenue* qui est celle que le ordinateur peut fournir en pratique dans un fonctionnement normal.

politique d'ouverture, en particulier vers le monde industriel (EDF, SNECMA, ONERA), qui le cofinance, et de favoriser les échanges scientifiques entre partenaires ; le CCRT dispose principalement d'une machine HP SC 45 de 2,4 teraflops et d'une machine NEC SX-6 de 0,4 teraflop.

- A cela s'ajoute TERATEC qui est une initiative du CEA pour développer un centre de compétence en calcul intensif et algorithmique qui permette à ses utilisateurs de se retrouver et d'échanger sur leurs approches (ingénieurs CEA, chercheurs universitaires, CNRS, mais aussi des industriels comme Bull, EDF ou SNECMA) ; TERATEC reste essentiellement une structure d'accueil du CEA, largement financée sur fonds propres.

Le département Sciences de la simulation et de l'information du CEA-DAM compte 190 personnes. Il reçoit 15 à 20 stagiaires et produit 5 ou 6 thèses par an. Il se positionne délibérément sur le créneau de la recherche technologique. Au fil des ans, il a développé un grand savoir-faire en ingénierie architecturale, indispensable à l'exercice de son rôle de maître d'ouvrage de grands systèmes, et en matière de logiciels, notamment de logiciels libres (*open source*) qu'il utilise largement, dans un souci d'indépendance technologique.

Le centre de calcul de l'IN2P3

Enfin, le centre de calcul (CC-IN2P3) de l'Institut National de Physique Nucléaire et de Physique des Particules (IN2P3) est l'une des 18 entités de l'IN2P3, lui-même laboratoire du CNRS. Le CC-IN2P3 a le statut d'unité de service du CNRS. Il a été créé pour répondre aux besoins des laboratoires travaillant dans les domaines de la physique des particules, de la physique des astroparticules et de la physique hadronique et de la matière nucléaire. Il a une approche client pour une quarantaine d'expériences internationales, pluriannuelles et impliquant des laboratoires de l'IN2P3. Le CC-IN2P3 travaille en fait depuis 50 ans dans un cadre européen, celui du CERN, et depuis une dizaine d'années dans un cadre mondial. En terme de calcul proprement dit, les besoins de l'IN2P3 se caractérisent essentiellement par la grande quantité de données à manipuler et à stocker (de l'ordre du petaoctet (10^{15} octets), dans un futur proche pour certaines expériences), alors que les calculs sont massivement parallèles. Le CC-IN2P3 semble bien répondre aux besoins spécifiques d'une communauté scientifique bien organisée, au sein de laquelle il a trouvé sa place tant au niveau européen qu'au niveau international et au travers de laquelle il s'est doté d'un savoir-faire important en ingénierie logicielle⁴.

Le CC-IN2P3 dispose de fermes de calculateurs (1,8 teraflop). Il emploie environ 60 personnes, dont 50 % à statut CNRS et 45 % de CDD.

La mise en service prochaine au CERN du projet de *Large Hadron Collider* (LHC), dont les résultats seront mis à la disposition de la communauté mondiale, a conduit le CC-IN2P3 à s'impliquer fortement dans les *grilles* de calculateurs (voir ci-dessous pp. 16 et 21) et à participer au programme européen EGEE, coordonné par le CERN.

⁴ Notamment au travers du logiciel ROOT d'exploitation de données expérimentales (<http://root.cern.ch/>) et de ses utilisations sur les expériences Virgo et GANIL.

Synthèse

En fin de compte, il apparaît que, sur les quatre grands centres de calcul existant en France, seuls le CINES et l'IDRIS constituent réellement des centres à vocation généraliste, ouverts à l'ensemble de la communauté scientifique. Le centre de calcul de l'IN2P3 (CC-IN2P3) est, lui, orienté très nettement projet autour de la physique des hautes énergies. Pour ce qui est du CEA, s'il est à l'évidence impossible de prendre en compte les moyens militaires de la DAM dans une problématique du calcul scientifique, il concourt clairement au travers du CCRT à la satisfaction de besoins scientifiques propres aussi bien qu'externes. En tout état de cause, parmi les quatre centres, c'est certainement celui du CEA qui constitue la meilleure référence en matière de gouvernance de moyens informatiques.

Les budgets des quatre centres sont difficiles à cerner avec précision, pour plusieurs ordres de raisons :

- les coûts de personnels n'apparaissent jamais dans les budgets ;
- les investissements ont connu ces dernières années des à-coups qui ôtent beaucoup de leur signification aux montants d'une année donnée ;
- les sources de financement sont multiples : ministère délégué à la Recherche en direct pour le CINES ou via le CNRS pour l'IDRIS et le CC-IN2P3, ministère de la Défense pour le CEA-DAM, ministère délégué à l'Industrie pour des opérations ponctuelles au CEA, contrats extérieurs de recherche (CC-IN2P3, qui les utilise pour recruter du personnel en CDD) ou prestations de services pour des industriels (CCRT), le tout complété pour certains de ressources propres⁵ ;
- enfin, pour des raisons compréhensibles, le CEA n'a pas communiqué de chiffres détaillés.

Sous ces réserves, et plutôt que de chercher à reconstituer avec précision le budget 2004 des quatre grands centres, le tableau suivant tente, à partir des données recueillies par la mission et au prix de comparaisons et d'extrapolations, de reconstituer les ordres de grandeur des financements publics, en régime permanent, pour les quatre grands centres nationaux de calcul intensif.

Centre	Fonctionnement ⁶ M€	Investissement M€	Personnel M€	Coût total annuel (M€)
CINES	1,7	2,8 ⁷	0,9	5,4
IDRIS	4,1	2,8 ⁸	2,3	9,2
CEA/CCRT	4,0 ⁹	6,0 ¹⁰	p.m.	10,0
CC-IN2P3	3,2	3,5	2,1	8,8
Total	13,0	15,1	> 5,3	33,4

⁵ Produits de placements financiers pour le CINES, contrats extérieurs pour le CC-IN2P3, financement par des industriels pour le CCRT.

⁶ A l'IDRIS, la rubrique fonctionnement comptabilise les investissements, hors grosses machines.

⁷ Pour obtenir un chiffre significatif en dépit des à-coups qu'ont connus les investissements du CINES ces dernières années, on a retenu en fait le montant annuel de l'amortissement linéaire sur 5 ans des matériels qui y sont actuellement en service.

⁸ Au cours des 5 dernières années (2000-2004), l'IDRIS n'a procédé à des investissements qu'en 2000 (5,7 M€) et en 2003 (6,2 M€). Pour obtenir un chiffre significatif, on a retenu en fait la moyenne des investissements effectués sur la période 1995-2004 (27,8 M€ au total).

⁹ Y compris les frais de personnel, mais hors frais de siège du CEA.

¹⁰ Moyenne des investissements effectués par le CEA au CCRT sur la période 2003-2005 (19 M€). Pour mémoire le CEA finance 72 % des coûts du CCRT.

C'est donc un peu plus de 35 M€ qui sont dépensés annuellement dans les quatre grands centres de calcul français, la moitié en investissements, un tiers en fonctionnement et un sixième en personnel.

III- Les mésocentres et les moyens des laboratoires

Ces dernières années ont vu l'apparition de centres de calcul de taille intermédiaire, dits *mésocentres*, au sein d'universités notamment. Ces centres sont destinés à mutualiser à un niveau régional les besoins de différents laboratoires ou d'équipes de recherche. Ils offrent des possibilités raisonnables de calcul de proximité à des équipes qui n'ont pas la possibilité de s'équiper individuellement. Les mésocentres permettent ainsi de traiter de grands volumes de calcul sans engorger les grands centres nationaux, dont l'utilisation ne se justifie que pour des calculs d'une autre ampleur. Leur gestion est aussi beaucoup plus souple que celle des grands centres.

Si l'utilité de ces mésocentres ne doit donc pas être remise en cause, il n'en va pas de même de leur mode de financement. Les rapporteurs ont tenté, avec l'appui de la direction de la Recherche, d'appréhender les budgets des mésocentres, dont une partie est financée par le ministère délégué à la Recherche et une autre par les régions, sans que cela épuise le sujet. De fait, les réponses obtenues se sont révélées difficilement exploitables. L'ampleur de la tâche justifierait en fait à elle seule une mission spécifique, indispensable pour avoir une vue exhaustive des moyens disponibles et des sommes qui y sont consacrées. En toute hypothèse, compte tenu des moyens limités dont il dispose, le ministère délégué à la Recherche doit se concentrer sur une stratégie d'équipement des grands centres nationaux. Sauf exception dûment justifiée, il doit donc laisser le financement des mésocentres aux universités, qui peuvent éventuellement y contribuer grâce aux fonds qu'elles obtiennent au titre du volet recherche du contrat quadriennal, et aux régions.

Enfin, certains laboratoires se sont équipés en moyens informatiques relativement lourds. Il s'agit le plus souvent d'initiatives locales de groupes d'utilisateurs très actifs qui peuvent ainsi gérer leurs calculs au plus près de machines, dont ils sont souvent allés rechercher le financement. Le côté dynamique d'une telle approche, qui peut être à l'origine du développement de certains logiciels ou de techniques de programmation innovantes, ne doit pas pour autant faire perdre de vue le nécessaire équilibre à respecter entre moyens individuels et moyens collectifs, notamment en termes de financement.

La hiérarchisation à trois niveaux (grands centres, mésocentres, moyens propres des laboratoires) du système français est saine dans son principe, mais on peut s'interroger sur le point d'équilibre à atteindre entre les trois catégories. En outre, selon la formule très parlante d'un interlocuteur de la mission, *le système français encourage les voyageurs sans billet*. Il est en effet facile de compter sur *les autres* pour obtenir les moyens informatiques dont on a besoin. La gratuité de l'accès aux grands moyens de calcul n'entre pas pour rien dans ce comportement.

La dernière conférence *Supercomputing 2004* tenue à Pittsburgh (Etats-Unis) a d'ailleurs souligné la mauvaise utilisation générale des grosses machines aux Etats-Unis¹¹, l'immense majorité des programmes y utilisant moins de 32 processeurs. Cela n'empêche au demeurant pas lesdites grosses machines d'être saturées par ces petits programmes. Plusieurs indices donnent à penser que la situation française n'est guère différente. Selon une analyse rapide effectuée par le CINES sur son ordinateur IBM Power 4 (256 processeurs) à la demande des rapporteurs :

- 10 % seulement des heures produites utilisent 128 processeurs ou plus (la moitié de la machine) ;
- 31 % seulement des heures produites utilisent 64 processeurs ou plus (le quart de la machine) ;
- 54 % seulement des heures produites utilisent 32 processeurs ou plus (le huitième de la machine).

Il serait utile à la réflexion globale de réaliser une étude plus détaillée sur ce sujet au CINES et à l'IDRIS. Il apparaît en tout cas d'ores et déjà que certains petits utilisateurs recourent aux calculateurs de l'IDRIS uniquement pour bénéficier d'un accès gratuit à ses bibliothèques de programmes, alors qu'ils n'y sont pas à leur place en termes de ressources nécessaires, uniquement pour ne pas avoir à payer de droits d'usage de logiciels.

IV- Le diagnostic

Le système français de calcul scientifique est en difficulté en raison de la conjonction de l'accroissement des besoins par le développement de nouvelles disciplines consommatrices de calcul et leur intensification par la substitution progressive de la simulation numérique à l'expérimentation traditionnelle et du début de la course à la puissance de la fin des années 90 dans laquelle la France ne s'est pas encore positionnée.

L'université de Mannheim et l'université du Tennessee se sont associées pour publier, deux fois par an, le *TOP 500*¹², classement des 500 calculateurs les plus puissants au monde. Cette liste, dont la version la plus récente date de novembre 2004, constitue une référence mondiale, même si elle présente trois défauts principaux reconnus. Le premier est que ce classement n'est pas exhaustif : certains utilisateurs, soit pour des raisons de confidentialité, soit qu'ils n'en voient pas l'intérêt, ne déclarent pas leurs machines. Le second est que le TOP 500 n'est qu'un instantané dans un monde en évolution permanente où la durée de vie d'une machine ne dépasse pas 5 ans et où, en pratique, elle fait l'objet d'augmentations constantes de configuration (*upgrades*). Son troisième défaut est qu'il s'appuie sur l'exécution de programmes de tests arbitraires (*Linpack*) et donc non nécessairement représentatifs de l'utilisation réelle des machines¹³. Cela dit, si le TOP 500 exerce à l'évidence une dictature sur les esprits, il n'en a pas moins le mérite d'exister, de permettre des comparaisons et, surtout, de permettre de mesurer les évolutions.

¹¹ Cette situation justifie l'existence de programmes incitatifs importants conduits par le Department of Energy (DOE), tel INCITE (<http://www.lbl.gov/CS/Archive/news073103.html>), dont les critères de sélection sont ainsi définis : *Successful INCITE proposals will describe high-impact scientific research and will be peer reviewed both in the area of research and also for general scientific review comparing them with proposals in other disciplines. Applicants must also present evidence that they can effectively use a major fraction of the 6,656 processors of the IBM SP supercomputer at the NERSC Center, which is the most powerful computer for unclassified research in the United States. Applicant codes must be demonstrably ready to run in a massively parallel manner on that computer.*

¹² <http://www.top500.org/>.

¹³ Selon certains interlocuteurs de la mission, certaines machines, comme la ferme de 2 200 processeurs Apple G5 du Virginia Polytechnic Institute auraient d'ailleurs été conçues avec le souci explicite d'optimiser le déroulement des programmes de test du TOP 500, de façon à figurer dans le haut du classement.

Dans le dernier classement (novembre 2004), la France ne place que 15 machines sur les 500 premières mondiales, la première étant 41^{ème} (CEA-DAM, 5,1 teraflops¹⁴). Les machines IBM de l'IDRIS n'apparaissent qu'au 175^{ème} rang (2,6 teraflop) et au 199^{ème} rang (2,6 teraflop). La machine vectorielle NEC SX-5 de l'IDRIS (214^{ème} en 2002) et les machines parallèles du CINES (349^{ème} en 2003) et de l'IN2P3 (167^{ème} en 2003) n'y figurent plus.

Si on regarde maintenant l'ensemble du classement :

- la première machine est américaine (Blue Gene, fruit d'une coopération entre IBM et le Department of Energy) avec une puissance de 70,7 teraflops, alors que la première machine française n'affiche que 5,1 teraflops (CEA-DAM) ;
- parmi les 20 premières machines mondiales, on en trouve 8 américaines, 2 japonaises (dont l'Earth Simulator, 3^{ème} au classement), 2 britanniques, une espagnole (centre de calcul de Barcelone, 2^{ème} au classement¹⁵) et une chinoise ;
- le Royaume-Uni compte 42 machines dans le TOP 500, dont 10 dans les 100 premières du classement¹⁶ ;
- l'Allemagne compte 35 machines dans le TOP 500, dont 5 dans les 100 premières du classement.

En matière de calcul scientifique, la puissance de la machine n'est évidemment pas tout, d'autant qu'elle reste difficile à mesurer dans l'absolu et qu'elle dépend largement de la façon dont on l'utilise. En outre, bien d'autres éléments entrent en ligne de compte dans la qualité de la recherche. C'est ainsi que l'Allemagne dispose de 2,5 fois la puissance de calcul de la France mais qu'il serait absurde d'en conclure que la science allemande est 2,5 fois meilleure que la française. Cela dit, il est inquiétant de constater que la diminution du nombre de machines française et le recul des machines restantes dans le classement s'accroissent depuis 2002. Alors que la France comptait 28 machines dans les 500 premières mondiales (5,6 %) en 2000, elle n'en a plus aujourd'hui que 15 (3,0 %). Alors qu'elle détenait 3,2 % de la puissance de calcul totale des 500 premières machines mondiales en 2000, elle n'en détient plus que 1,9 % aujourd'hui. Alors qu'en termes de puissance de calcul elle était 4^{ème} mondiale en 2002 et 5^{ème} en 2003, elle n'est plus que 6^{ème} en 2004. Les moyens français ont été bons jusque vers 2000. C'est la mise en service du supercalculateur japonais Earth Simulator qui a mis en évidence le retard français. Une simulation qui prend deux mois à l'IDRIS ne prend que quatre ou cinq jours sur l'Earth Simulator.

Or, qu'on le veuille ou non, l'environnement international du monde de la recherche se caractérise par une course à la puissance de calcul menée par le Japon et les Etats-Unis, mais dans laquelle s'inscrivent aussi nos voisins et partenaires, qui maintiennent ainsi leurs positions, alors que nous décrochons. De manière générale, les gouvernements américain, japonais, allemand et britannique sont proactifs en matière de calcul scientifique. Au-delà des seuls équipements, ces politiques ont un effet positif sur l'organisation des équipes de recherche et leur motivation à investir dans le calcul intensif.

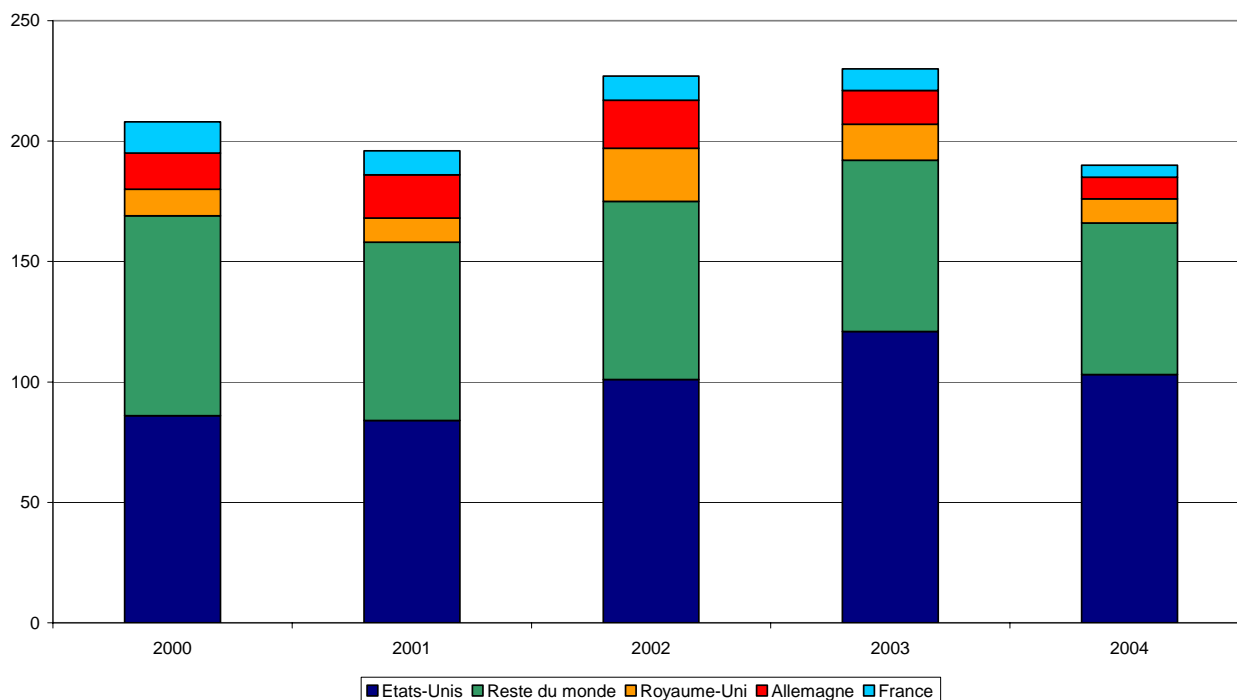
¹⁴ Cette machine, installée en 2001, doit être remplacée fin 2005 par une machine Bull NovaScale à 8 704 processeurs d'une puissance de 10 teraflops soutenus, soit de l'ordre de 60 teraflops crête. Toutes choses égales d'ailleurs, cette nouvelle machine se placerait au 2^{ème} ou au 3^{ème} rang du TOP 500 de novembre 2004. Ceci montre bien le caractère éminemment fugace de ce classement.

¹⁵ Le gouvernement espagnol a mis en œuvre une politique extrêmement volontariste en matière de calcul scientifique, avec l'installation à Barcelone d'un ordinateur IBM de 40 teraflops, connu sous le nom de *Mare Nostrum* et destiné aux applications médicales, aux études climatologiques et à la recherche sur les nouveaux matériaux, pour un investissement de 70 M€ Reste à voir ce que donnera sur le long terme cette politique qui a temporairement hissé l'Espagne à la deuxième place du TOP 500.

¹⁶ Conséquence de l'importance de la place de Londres dans le monde des affaires, 29 machines rentrent dans la catégorie *Industry* qui, ici, recouvre essentiellement des banques. On y trouve aussi, dans le domaine pétrolier, trois machines de la Compagnie Générale de Géophysique, société française qui exploite des forages pétroliers en mer du Nord.

Si on extrait maintenant du TOP 500 les seuls calculateurs catalogués *academic* ou *research*¹⁷, soit un parc de 190 à 230 machines selon les années, pour obtenir des références plus pertinentes pour ce rapport, la France qui y figurait pour 13 en 2000 n'y figurait plus que pour 9 en 2003 et pour 5 en 2004. Dans le même temps, elle passait de 2,3 % de la puissance de calcul totale à seulement 1,3 %. Il est vrai que la montée en puissance des pays asiatiques, Japon, Corée et maintenant Chine, tend mécaniquement à réduire la part des pays occidentaux. Il est vrai aussi que l'Allemagne semble faiblir, passant entre 2000 et 2005 de 15 machines à 9 et de 7,1 % de la puissance de calcul totale à 2,9 %. Quant à l'apparente progression britannique, elle demande à être nuancée par le fait que la moitié de sa puissance de calcul (33 teraflops sur 65) est le fait des deux machines du *European Center for Medium term Weather Forecast* (ECMWF) de Reading, organisation internationale, donc non vraiment britanniques. En revanche, la progression américaine suit assez rigoureusement la loi de Moore¹⁸. Il en va au demeurant de même pour les pays classés *reste du monde* dont les moyens de calcul sont à plus de 50 % ceux du Japon, de la Corée, de Taiwan, de Hong Kong et de la Chine et qui tendent à se rapprocher globalement des Etats-Unis.

nombre de calculateurs dans le TOP 500 "research" + "academic"



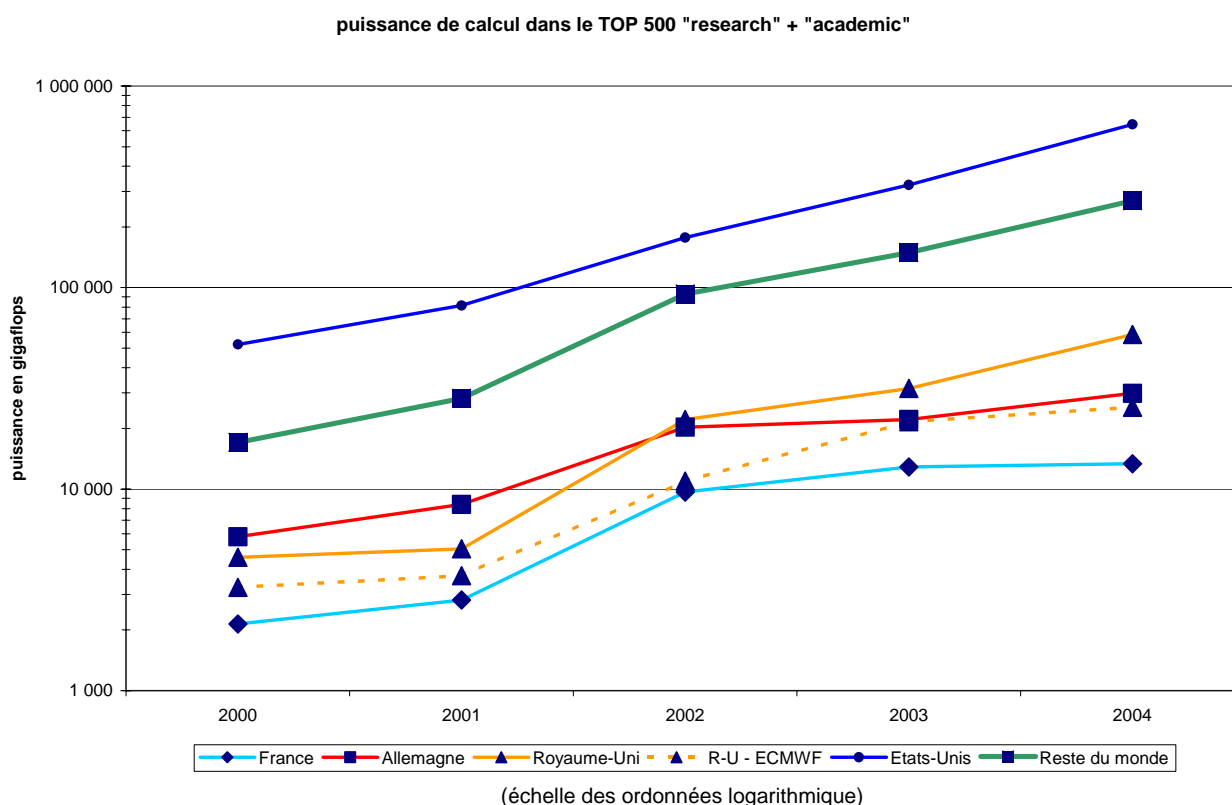
Cet état de choses n'aide évidemment pas nos chercheurs à bien se placer dans la compétition internationale, face à des équipes étrangères qui disposent de ressources supérieures. Cela obère leur capacité à participer en bonne place à des projets internationaux et européens et réduit encore un peu plus l'attrait de la recherche française. A l'inverse, il n'est pas indifférent de savoir que

¹⁷ Sont ainsi écartées les machines étiquetées *classified*, *government* et *industry* (au sens anglais du terme, c'est à dire non seulement l'industrie au sens manufacturier du terme, mais aussi les activités de services, comme les banques, les compagnies d'assurance, les opérateurs de télécommunications, etc.).

¹⁸ Loi empirique qui affirme que la puissance des calculateurs double environ tous les 18 mois à 2 ans. De fait, la puissance de calcul américaine a été multipliée par 12 entre 2000 et 2004, soit un doublement tous les 22 mois, alors que la puissance française n'a été multipliée que par 6 pendant le même laps de temps.

le Japon paye les frais de premier séjour de chercheurs européens pour qu'ils viennent utiliser gratuitement sur place l'Earth Simulator. Cette démarche est rien moins qu'innocente car elle vise bien à terme à assécher la recherche européenne, dans un domaine, la climatologie, où le Japon ne se situe pas aujourd'hui dans les meilleurs mondiaux. En encourageant par défaut de tels errements, on contribue de fait à délocaliser l'expertise scientifique nationale et à la transférer à d'autres.

Les statistiques publiées par l'organisation ARCADE¹⁹, qui ramènent l'effort des principaux pays européens en matière de calcul intensif à la taille de chacun d'eux, exprimée en termes de population ou de produit intérieur brut, donnent également une image attristante de la situation, dans laquelle notre pays apparaît systématiquement en avant-dernière position, loin derrière le Royaume-Uni ou l'Allemagne, mais aussi très loin derrière le Danemark ou la Suède.



Le CINES souffre à l'évidence de sa petite taille, tant vis-à-vis d'autres centres européens avec lesquels il pourrait envisager des partenariats, Barcelone notamment (40 teraflops), que vis-à-vis de l'IDRIS, lui-même à la marge par rapport à ses partenaires de DEISA²⁰. En outre, le statut d'EPA n'apparaît pas particulièrement bien adapté à sa situation. Enfin, pour universitaire qu'il soit, l'environnement montpelliérain ne constitue pas un terrain de choix pour le calcul intensif.

Quant au CNRS, avec l'IDRIS il se retrouve responsable d'une mission nationale de calcul scientifique qu'il n'a pas vraiment demandée et dont on ne l'a pas vraiment chargé. Le rattachement de l'IDRIS au département STIC du CNRS n'a d'ailleurs pas de sens. A tout le moins, s'il

¹⁹ <http://www.arcade-eu.info/index.html>. ARCADE est une organisation européenne non-gouvernementale à but non lucratif destinée à favoriser les échanges de savoir et d'informations sur les politiques et les infrastructures en matière de calcul intensif.

²⁰ Sur DEISA, voir ci-dessous p. 15.

devait rester dans le cadre du CNRS, il conviendrait de changer son statut d'unité propre du département STIC en celui d'unité mixte de service (UMS).

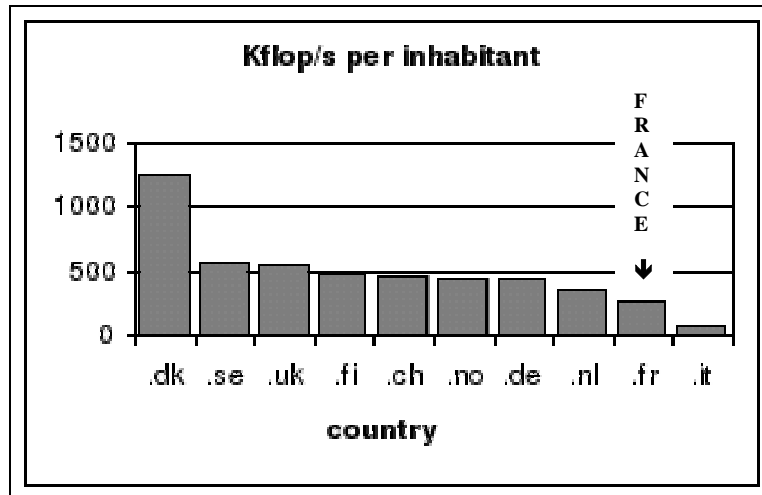


Figure 1 : puissance de calcul par habitant en Europe

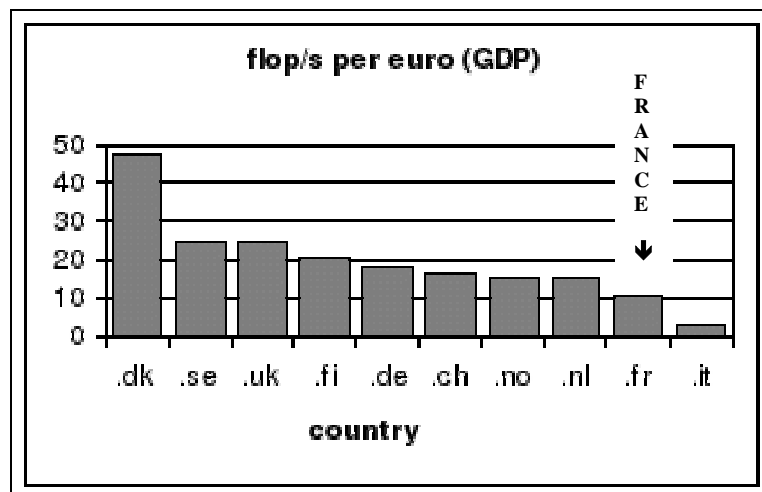


Figure 2 : puissance de calcul rapportée au PIB en Europe

Cela dit, même s'ils ont des statuts différents, des tutelles différentes et des gestions distinctes, le CINES et l'IDRIS ont une approche coordonnée face à la communauté scientifique, grâce à l'existence de comités thématiques communs. Au nombre de neuf²¹, ces comités thématiques couvrent les différents domaines des sciences qui utilisent le calcul intensif. Leur rôle consiste à évaluer les demandes de ressources informatiques des laboratoires et à les présenter aux deux conseils scientifiques du CINES et de l'IDRIS.

Cette approche commune sera complétée en mars 2005 par la mise en réseau des calculateurs du CINES et de l'IDRIS, qui fera que l'utilisateur ne les verra plus que comme un centre unique.

Pour le reste, le CINES et l'IDRIS se comportent essentiellement comme de simples prestataires de service, sans vraie valeur ajoutée complémentaire et sans activité de recherche propre. Ce choix, qu'il soit délibéré ou subi, leur est nuisible, tant il est vrai que les grands centres de calcul étrangers entretiennent autour d'eux une importante activité de recherche dans le domaine du calcul scientifique. Ils peuvent ainsi animer efficacement une activité de service aux utilisateurs scientifiques et les soutenir dans la mise en œuvre et la conception de leurs codes²². Le problème français semble donc largement structurel.

Le système est d'autant plus sclérosé que la rotation du personnel est faible. Elle est même carrément nulle au CINES. La moyenne d'âge y est de 50 ans et conduira à des départs massifs à la retraite entre 2007 et 2014, sans qu'on voie de piste sérieuse pour assurer le renouvellement du personnel. Au demeurant, le CINES et l'IDRIS se heurtent à des difficultés de recrutement du fait des niveaux de salaires qu'ils peuvent proposer dans un univers hautement concurrentiel, même s'ils bénéficient de l'effet de carte de visite de l'Université qu'ils procurent.

Au vrai, un certain nombre de communautés scientifiques françaises semblent manquer d'ambition en matière de calcul et se réfugier commodément derrière la persistance des contraintes budgétaires pour justifier l'absence de leur part de toute réflexion prospective et stratégique en la matière.

Des quatre grands centres français, le CEA apparaît finalement comme le seul à avoir une vraie gouvernance de ses moyens informatiques. Le CINES et l'IDRIS fonctionnent largement sur la base de la reconduction du passé, quand ils ne mènent pas leur propre politique. La transparence et la prise en compte des priorités gouvernementales ne semblent pas les qualités premières du système²³. L'organisation en comités thématiques tend d'ailleurs à conforter les situations acquises (climatologie, sciences de l'Univers, ...) et à ignorer les besoins émergents (biologie, matériaux, nanotechnologies, ...). De fait, le taux de renouvellement des utilisateurs ne dépasse pas 20 % par an. Quant à l'effort de prospective, il n'est bien développé qu'au CEA et à Météo -France, tous deux soumis à des contraintes fortes, de caractère militaire pour le premier, de caractère opérationnel pour le second.

Pour le reste, la réflexion sur les besoins à moyen ou long terme des différentes communautés scientifiques reste extrêmement limité en dehors de l'effort ponctuel effectué à la

²¹ Environnement ; mécanique des fluides ; milieux réactifs ; astrophysique, géophysique, terre solide ; électromagnétisme et plasmas chauds ; mathématiques, mathématiques appliquées, systèmes modèles ; systèmes moléculaires organisés et biologie ; chimie quantique et modélisation moléculaire ; physique, chimie et propriété des matériaux.

²² Comme les services d'assistance et d'optimisation fournis par le Computer Services for Academic Research (CSAR) britannique (<http://www.csar.cfs.ac.uk/services/support/index.shtml>) ou ceux fournis par le CERN à ses utilisateurs (<http://services.web.cern.ch/services/>).

²³ C'est ainsi que le GIP Mercator, qui regroupe le CNES, le CNRS, l'IRD, l'IFREMER, Météo France et le SHOM, n'a pu obtenir pour l'année 2005 que 150 000 heures de calcul pour 410 000 heures demandées sur le calculateur IBM de l'IDRIS au titre du projet franco-allemand d'océanographie Drakkar. Dans le même temps la direction du centre a gelé des heures sur cette machine en vue d'une utilisation dans le cadre de DEISA.

demande des rapporteurs²⁴, alors qu'il pourrait utilement éclairer aujourd'hui les futurs choix de machines et l'orientation des travaux de développement de codes appelés à fonctionner sur d'autres générations de machines que celles d'aujourd'hui.

La multiplicité des sources de financement du calcul haute performance (Education nationale, CNRS, Défense, Industrie, Commission européenne, contrats extérieurs, financements industriels, ressources propres, prêts de personnel, ...) et la multiplicité des statuts juridiques (établissement public industriel et commercial pour le CEA, établissement public administratif pour le CINES, laboratoires du CNRS pour l'IDRIS et le CC-IN2P3), sans parler de celle des statuts des personnels, ajoutent à l'opacité du système.

Le CINES et l'IDRIS n'ont pas de système de facturation, même *pro forma*. Il est frappant de voir qu'on ne connaît pas le coût réel de l'heure de calcul en France. Malgré des recommandations anciennes sur les grands équipements, on se trouve face à une économie non monétaire, où ne sont jamais comptés les coûts d'un personnel qui relève de statuts et d'organismes divers. Ceci rend impossible toute comparaison sérieuse avec des solutions d'infogérance. Cela n'incite pas non plus les utilisateurs à calibrer leurs demandes au plus juste.

Au final, il est frappant de constater qu'en dehors du CEA-DAM, dont les exigences de la dissuasion nucléaire garantissent la pérennité des moyens de calcul, l'équipement des autres centres (CINES, IDRIS, IN2P3) a procédé sans planification sérieuse à moyen ou long terme et sans vraie continuité dans les budgets²⁵, dans un domaine où les machines sont obsolètes en 3 ou 4 ans, alors même que le développement des grands codes demande 5 ans de travail et restent utilisés pendant plus de 10 ans.

Si cette remarque est vraie pour chacun de ces trois centres pris individuellement, elle l'est a fortiori au niveau global. On ne voit pas en France d'instance organisée où se discutent les choix stratégiques à long terme en matière de calcul intensif. On voit encore moins d'instance qui appréhende le problème dans sa globalité et pas uniquement sous l'angle de l'augmentation de puissance des machines, mais qui tienne compte aussi des évolutions en matière d'ingénierie logicielle, de nouvelles applications (apparition de la simulation numérique, de la restitution graphique des résultats, ...), voire de l'émergence de nouveaux domaines d'utilisation tels que ceux déjà cités (biologie, nanotechnologies, ...). Dans un secteur d'activité en évolution aussi rapide, il est pourtant indispensable d'être à l'écoute de ce qui se passe dans le monde, de ne pas céder aux effets de mode, voire de ne pas tomber dans le piège d'annonces destinées uniquement à leurrer ou à induire en erreur des concurrents étrangers. Il est donc essentiel d'être vigilant, critique et réactif si quelque chose de nouveau se passe soudainement. En un mot, le développement du calcul scientifique en France ne peut faire l'économie d'une vraie démarche d'intelligence économique.

²⁴ C'est une raison supplémentaire pour saluer ici les initiateurs de l'atelier tenu au CERFACS (voir p. 19) le 1^{er} septembre 2004 sur le thème *Le climat du XXI^{ème} siècle et les simulations frontière*, ainsi que la direction de l'INSU qui a mis en place, pour la deuxième fois, un Groupe d'Etudes Prospectives sur le Calcul Intensif (GEPCI 2). Les travaux du CERFACS font ressortir que, dans la limite du thème retenu, les scientifiques ne pourront apporter aux politiques de réponses à un certain nombre de questions déterminantes pour l'avenir de l'humanité que si les puissances de calcul actuelles sont multipliées par un facteur de 10^2 à 10^5 à horizon de 10 à 20 ans, selon les thématiques. Ces chiffres n'ont au demeurant rien d'absurde puisque, selon la loi de Moore déjà évoquée, une croissance d'un facteur 10^4 est atteinte en 20 ans. Quant au rapport du GEPCI 2, il fait le point des évolutions attendues en matière de programmes de climatologie, d'océanographie et de météorologie sur la période 2005-2008.

²⁵ C'est l'occasion de souligner que le statut d'EPA du CINES et l'obligation qu'il lui crée d'avoir des comptes propres et de pratiquer des amortissements comptables lui assure, année après année, la constitution d'une réserve financière qui facilite grandement, le moment venu, le financement de machines de remplacement.

V- La France et l'Europe

Si la France se doit de combler son retard en matière de calcul scientifique et de mettre en œuvre une programmation nationale, elle ne peut pour autant éviter de passer à la dimension européenne si elle veut prétendre se hisser aux niveaux du Japon et des Etats-Unis. Au demeurant, dans ce contexte, l'adjectif *européen* ne vise pas nécessairement le seul cadre de l'Union européenne mais aussi bien des coopérations multilatérales. La thématique calcul scientifique en tant que telle n'existe d'ailleurs pas à Bruxelles.

La dimension européenne du calcul intensif est multiforme. Elle concerne :

- l'accord DEISA, promu par l'IDRIS, qui vise à échanger de la puissance de calcul mais qui, s'il limite la course à la puissance, ne la supprime pas complètement dans la mesure où chaque partenaire, tout à la fois pour des raisons de prestige et de pouvoir, cherche à préserver son poids relatif dans l'ensemble ;
- le projet EGEE, porté par le CERN et relayé en France par le CC-IN2P3, dans lequel l'internationalisation est consubstantielle à la nature même des projets de grilles de PC ;
- la préparation du 7^{ème} PCRD (Programme Commun de Recherche et de Développement) ;
- enfin, le projet européen d'infrastructure de calcul intensif en cours de négociation entre l'Allemagne, la France et le Royaume-Uni.

DEISA

Le projet DEISA (*Distributed European Infrastructure for Supercomputing Applications*) vise à créer, par interconnexion des moyens existants, un centre de calcul virtuel entre la France, l'Allemagne et l'Italie, avec un élargissement prévu à 11 partenaires, hors Royaume-Uni. Pour intéressante qu'elle soit tant du point de vue de l'europanisation que de l'optimisation de l'utilisation des machines, l'approche DEISA ne résoudra pas les problèmes de gros calculs couplés. Elle reste bornée par le fait qu'elle ne fournira jamais l'équivalent d'un supercalculateur comparable aux meilleures machines disponibles au Japon ou aux Etats-Unis. DEISA ne peut de fait offrir de puissance de calcul supérieure à celle qu'offre le calculateur le plus puissant du réseau. En revanche, DEISA permet une optimisation de la répartition des charges entre calculateurs. Ainsi, l'IDRIS n'affecte actuellement pas plus du quart d'une machine à une seule tâche. Avec DEISA, il sera possible d'envoyer des jobs banalisés en Allemagne ou en Italie et de libérer ainsi les 3/4 de la machine pour un seul utilisateur. Il faut signaler que dans DEISA les droits de vote dépendent de la puissance installée. Si la personnalité du directeur de l'IDRIS et promoteur du projet a permis jusqu'à présent à la France de jouer un rôle moteur dans ce projet, sa position pourrait s'effriter rapidement en l'absence de renouvellement de nos grands moyens de calcul.

En revanche, la démarche DEISA prendrait tout son sens dans le cadre du projet européen d'infrastructure de calcul intensif, détaillé ci-après (voir pp. 16 et 27), dans lequel elle devrait s'inscrire.

Les grilles

Pour sa part, le CC-IN2P3 participe aux projets européens de grilles²⁶ de calculateurs Datagrid et EGEE, avec l'objectif d'avoir une capacité de production académique et industrielle. Le développement des grilles présente un intérêt tout particulier pour le dépouillement des futures expériences menées au CERN sur le LHC dans le cadre du programme LCG (*LHC Computing Grid*), mais aussi potentiellement pour la biologie, les sciences de la Terre et, pourquoi pas, à terme pour la climatologie. Datagrid et EGEE (*Enabling Grids for E-science in Europe*) sont des projets de la Commission européenne. L'objectif de Datagrid est de bâtir une infrastructure de prochaine génération de calcul intensif et de partage de bases de données à grande échelle. EGEE de son côté rassemble les experts de 27 pays pour développer en Europe une infrastructure de grille de services accessibles en permanence aux scientifiques.

Le 7^{ème} PCRD

Le 7^{ème} PCRD européen est en cours d'élaboration. Le moment est donc opportun pour y faire prendre en compte les besoins du calcul scientifique. On peut se demander s'il ne serait pas intéressant d'y inclure le développement d'un calculateur de hautes performances qui, compte tenu de son coût, ne peut être développé qu'en coopération européenne et dont le marché est trop étroit pour permettre un financement purement industriel. L'ambition ne serait alors pas tant de se livrer à une course à la puissance avec les Etats-Unis ou le Japon que d'assurer à l'Europe le maintien de compétences en matériel et en logiciel, mais aussi une certaine indépendance technologique garante de son indépendance politique.

A cet égard l'annonce récente par le CEA du choix d'une machine Novascale de Bull pour son programme TERA 10 est peut-être un signe. Un regroupement au niveau européen des moyens de Bull avec ceux de Quadrics (Royaume-Uni²⁷) et de Scali (Norvège) en matière de réseaux, ainsi qu'avec ceux de PME allemandes du logiciel pourrait faire sens. En toute hypothèse, la préparation du 7^{ème} PCRD fournira à l'évidence une occasion de se prononcer sur ce type d'objectif.

L'initiative tripartite Allemagne - France - Royaume-Uni

Le projet d'infrastructure européenne de calcul intensif en cours de négociation entre l'Allemagne, la France et le Royaume-Uni paraît extrêmement intéressant, dans la mesure où ce n'est que par une approche et des financements multilatéraux qu'il sera possible de se procurer et d'exploiter des machines qui se situent dans les toutes premières mondiales par leur puissance. L'idée est que les trois pays achètent tous les deux ans un calculateur de très hautes performances qui serait installé à tour de rôle dans chacun des trois pays. Un pays donné verrait ainsi sa machine renouvelée tous les 6 ans, ce qui correspond *grosso modo* à la durée de vie normale de ce type de matériel. Selon

²⁶ Les termes *grille*, en français, ou *grid*, en anglais, désignent des réseaux d'ordinateurs qui peuvent se partager l'exécution d'une même tâche. En pratique, ils recouvrent des réalités extrêmement diverses qui vont de l'interconnexion de quelques grands calculateurs (DEISA) à celle de milliers d'ordinateurs personnels (*Personal Computers* ou PC), soit au travers de réseaux dédiés (RENATER, GEANT), soit par Internet.

²⁷ Quadrics est une société britannique, mais son capital est détenu par la holding publique italienne Finmeccanica.

les meilleures estimations actuelles, l'investissement se monterait à 200 M€ tous les deux ans et le fonctionnement à 20 M€/an les deux premières années, 40 M€/an les deux années suivantes, pour se stabiliser à 60 M€/an au-delà. Dans l'hypothèse d'un soutien à hauteur de 35 % de la Commission européenne et de 15 % d'autres entités, dont des industriels, ainsi que de l'utilisation du système par d'autres pays associés, le coût moyen annuel pour la France serait de 18 M€, ce qui paraît raisonnable au regard des enjeux.

La hiérarchie des centres de calcul pourrait ainsi se voir adjoindre un quatrième niveau, européen, dédié aux véritables très gros programmes, à qui ne suffiraient pas les trois niveaux nationaux. Comme indiqué plus haut, dans une telle perspective, l'initiative DEISA prendrait un sens nouveau, dans la mesure où elle préfigurerait l'interconnexion des trois centres de grande puissance installés dans chacun des trois pays.

Le rapport annexe, réalisé sur la base des contributions reçues par les rapporteurs, donne des indications sur les principales avancées scientifiques que rendrait possible cet accroissement significatif des moyens de calcul à la disposition de nos chercheurs.

VI- Les logiciels

Le rôle des logiciels dans la mise en œuvre des grands calculateurs scientifiques est fondamental. Il se situe à trois niveaux :

- les systèmes d'exploitation qui permettent au calculateur de fonctionner et à l'utilisateur d'avoir accès aux services de base ;
- les compilateurs qui assurent la traduction du programme en langage machine, son implantation et son exécution ;
- les programmes d'application ou logiciels applicatifs qui permettent de résoudre le problème spécifique du chercheur (climatologie chimie quantique, ...).

Les couches d'*intergiciel* et de *run time* sont spécifiques du calcul scientifique (parallélisme dans les applications, systèmes de fichiers parallèles de grande taille, optimisation de la performance, fiabilisation du fonctionnement d'ensemble face à la défaillance de composants²⁸). Elles exigent une connaissance intime de la machine. Elles recourent très largement aux logiciels libres (MPI, systèmes de fichiers parallèles, ...). Ceci résulte en partie de l'étroitesse du marché, mais aussi de l'impulsion constante donnée par les grands utilisateurs publics. Toutefois l'optimisation et l'*industrialisation* (fiabilisation, robustesse, documentation, ...) de ces logiciels sont souvent réalisées avec l'aide d'équipes industrielles financées par les constructeurs.

Les logiciels système exigent une parfaite intégration avec le matériel, pour assurer le niveau requis de fiabilité. Pour autant, le logiciel libre Linux tend à se généraliser, à côté de logiciels très spécifiques réalisés en collaboration étroite entre certains grands utilisateurs et les constructeurs d'ordinateurs, qui s'entourent parfois d'une communauté travaillant selon le principe du logiciel libre.

Les compilateurs constituent un quasi-monopole des constructeurs américains et japonais d'ordinateurs (IBM, SGI, ...) et des fabricants de microprocesseurs (Intel, ...). En créant une

²⁸ Ce problème devient critique avec les architectures actuelles de calculateurs qui regroupent des milliers de processeurs. Si un processeur a un MTBF (*Mean Time Between Failure* ou temps moyen entre pannes) de 10^{-8} s^{-1} , c'est-à-dire environ une panne tous les 3 ans, un calculateur à 1 000 processeurs connaîtra une panne de processeur toutes les 100 000 secondes, soit toutes les 28 heures. L'exécution d'un grand code qui nécessite plusieurs jours de calcul aura donc de fortes chances de se trouver confronté à une panne de processeur...

concurrence, l'apparition de la suite de logiciels libres GCC de la FSF a contribué à l'amélioration des performances et de la qualité mais aussi à l'extrême concentration du marché. C'est un bon exemple d'une communauté du logiciel libre qui a pu se créer et influencer la situation concurrentielle du marché, en mettant en œuvre les résultats scientifiques publiés et qui bénéficie de la coopération de nombreux acteurs industriels.

A ce niveau des systèmes d'exploitation et des compilateurs, les besoins techniques ne correspondent pas au marché visé par les constructeurs d'ordinateurs. Ceci conduit les grands centres de calcul, en particulier américains, d'une part, à préserver un savoir-faire important dans l'intégration de ces systèmes et, d'autre part, à s'appuyer sur une utilisation habile de développements en logiciels libres qui leur permettent d'intervenir à côté de constructeurs et d'intégrateurs. En France, le CEA-DAM conduit une série de projets de ce type pour s'assurer la maîtrise de logiciels clés, dans une perspective d'indépendance technologique.

A l'exception de rares secteurs comme la chimie quantique qui utilisent des codes commerciaux, les chercheurs développent les leurs. Cela correspond d'ailleurs bien à leur état d'esprit. Les équipes de recherche souhaitent conserver la maîtrise complète de la conception et du développement de leurs codes, ce qui se comprend, et ne voient les centres de calcul que comme des prestataires de service, ce qui se comprend moins. Les grands codes représentent le travail d'un groupe de chercheurs, pendant une période qui peut atteindre 5 ans et plus. Ainsi l'investissement dans la modélisation numérique ne peut se faire que dans l'optique de plans à moyen terme et les retards sont difficiles à combler.

Dans les domaines scientifiques où la simulation numérique est la mieux implantée, des méthodologies ont été mises en place pour garantir la bonne fin de la réalisation des codes et la possibilité de les exploiter suffisamment longtemps. La complexité croissante de ces développements devrait inciter à rapprocher le développement de ces grands logiciels de l'expertise en génie logiciel, ainsi qu'à une meilleure reconnaissance de ce type de travail dans la carrière des chercheurs.

Pour autant, beaucoup de secteurs scientifiques en France manquent de maturité en ingénierie des grands codes et ne disposent généralement pas des compétences informatiques nécessaires à leur optimisation et à leur industrialisation en vue de leur exploitation sur de très grosses machines et de leur portabilité sur d'autres machines que celles habituellement utilisées. Un chercheur seul, voire une équipe de recherche, ne peut plus développer ses codes par ses propres moyens. Il ne suffit pas d'intégrer quelques logiciels du commerce (simulation, visualisation, ...) à des éléments développés en interne. C'est pourquoi les équipes de recherche réclament des postes d'ingénieurs dédiés, mais n'envisagent ni mutualisation ni externalisation, dont elles ne comprennent pas vraiment l'intérêt. Il est significatif à cet égard que le CEA fasse un très gros effort sur l'ingénierie des codes. On ne peut d'ailleurs que déplorer l'absence de formation générale au calcul scientifique dans le système universitaire français.

Ces communautés communiquent certes sur le plan de l'architecture et de la meilleure utilisation des machines par le biais des comités d'utilisateurs mis en place par le CINES et l'IDRIS²⁹. En revanche, elles ne communiquent pas assez sur le plan des méthodes et des algorithmes, elles n'utilisent sans doute pas assez non plus les diverses bibliothèques qui existent en libre accès et ne participent du coup pas assez à leur développement. Le manque d'interdisciplinarité entre chercheurs, mathématiciens appliqués et informaticiens a frappé les rapporteurs.

A la notable exception près du CERFACS (une centaine de personnes) et de l'INRIA dans certains de ses projets, la France ne semble pas disposer d'organisations dédiées à ce que les Américains appellent les technologies algorithmiques et qu'ils estiment cruciales pour les progrès futurs du calcul intensif.

²⁹ Des structures similaires existent, avec un rôle encore plus important, pour la partie civile du CEA et à l'IN2P3.

Créé en 1987, le CERFACS (Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique) a pris en 1996 sa forme actuelle d'une société civile³⁰ dont les actionnaires sont EADS, SNECMA, EDF, le CNES et Météo France. Il s'est attaché au développement de la modélisation et à la formation des *computer scientists* pour la préparation des futurs outils de leurs actionnaires. Il reste donc dans le domaine de la mécanique des fluides au sens large. Il compte une centaine de personnes, chercheurs, ingénieurs, informaticiens et dispose d'un budget annuel de 6 M€ (salaires compris) provenant à 40 % de subvention de ses actionnaires et à 60 % de contrats externes. Il ne dispose pas de moyens de calcul propres importants, ce qui n'est d'ailleurs pas son objectif. Le positionnement du CERFACS est original dans le paysage français, alors qu'il est classique aux Etats-Unis ou en Allemagne (Fraunhofer). Le CERFACS essaye de nouer des relations avec le CINES et d'y placer ses produits.

La mission n'a en tout cas pas identifié de source de financement en matière d'ingénierie informatique pour le calcul scientifique en France. Au vrai, il ne semble pas exister en France de liaison entre les grands centres de calcul et le monde de la recherche en informatique, INRIA notamment. Pourtant le développement de codes optimisés, robustes, facilement portables et bien documentés procure une vraie valeur ajoutée qui résulte des gains de temps dans le développement des programmes, dans l'utilisation de la puissance de calcul et dans la transmission organisée du savoir.

A cet égard, il est intéressant de voir que le programme américain ASCI (*Accelerated Strategic Computing Initiative*) initié par le Department of Energy pour promouvoir la simulation numérique d'explosions atomiques comprend une forte composante informatique et calcul scientifique (*computer science*) qui confortera encore l'avance américaine en ce domaine. Le programme SciDAC³¹ destiné à satisfaire cet objectif dispose d'un financement de 60 M\$³². Il est tout aussi intéressant de voir que chez nous ce point n'a pas échappé à de bons esprits puisque le fait est rapporté dans le *Rapport scientifique 2003* de l'IDRIS³³. Ce thème de la nécessaire coopération entre chercheurs, spécialistes des mathématiques appliquées et *computer scientists*, de la création d'équipes multidisciplinaires et de la nécessaire formation d'informaticiens de haut niveau, est d'ailleurs lourdement récurrent dans les nombreux rapports américains récents consacrés au calcul intensif et qui y voient une des conditions d'un *leadership* des Etats-Unis en calcul de hautes performances³⁴.

En fait, c'est une approche nouvelle du problème qu'il faut envisager en recourant à un fonctionnement en équipe projet regroupant l'ensemble des compétences nécessaires, qu'elles soient scientifiques (climatologue, chimiste, biologiste, ...) ou informatiques (architecte, algorithmicien, programmeur, ...). Avec de gros moyens de calcul comme ceux d'aujourd'hui, il est indispensable de maîtriser la conception du logiciel. Un certain nombre de personnes rencontrées par la mission s'accordent d'ailleurs à dire que, rien qu'en soignant la conception d'un programme, il est possible de gagner un facteur 10 en efficacité dans son exécution. C'est d'ailleurs pourquoi le monde industriel qui développe de grands codes (télécommunications, aéronautique, défense, ...) recourt à des ateliers logiciels.

Une politique active pour participer aux développements des logiciels essentiels dans le cadre *open source* et pour rechercher le leadership sur certains des projets devrait conforter notre capacité à mettre en œuvre et intégrer des très grands systèmes. Elle permettrait également de soutenir l'offre industrielle, de matériels et de services. C'est la politique que met en œuvre le CEA autour de son projet TERATEC de pôle de compétences sur le calcul de hautes performances, en rassemblant

³⁰ Les ressources du CERFACS proviennent à 40 % de subventions de ses actionnaires et à 60 % de la vente de ses prestations à des tiers.

³¹ *Scientific Discovery through Advanced Computing* (<http://www.scidac.org/>).

³² Investissement pluriannuel (cf. http://www.sc.doe.gov/ascr/20040405_SciDAC_presentation.ppt).

³³ IDRIS, *Rapport scientifique 2003*, pp. 109 et 110.

³⁴ Notamment Office of Science, U.S. Department of Energy, *A science-based case for large-scale simulation*, 30 juillet 2003, disponible à l'adresse <http://www.pnl.gov/scales/>, et National Research Council of the National Academies, *Getting up to speed, the future of supercomputing*, 2004, disponible à l'adresse <http://www.nap.edu/books/0309095026/html/>.

scientifiques et industriels. La charte commune *logiciel libre* que viennent de signer l'INRIA, le CEA et le CNRS va dans le même sens. Il est possible d'intensifier les actions à partir de ces initiatives.

Il y a également nécessité absolue à dynamiser très fortement la recherche en logiciels en France et à travailler sur l'algorithmique. La solution pourrait consister à créer des structures analogues au CERFACS pour chacune des grandes thématiques de recherche.

Il paraît également souhaitable de prendre des mesures incitant les acteurs de la recherche publique à coopérer plus intensément avec les industriels français et européens (Bull, Scali) autour de plates-formes de calcul intensif.

VII- Besoins scientifiques et besoins industriels

Le calcul de haute performance est aujourd'hui largement utilisé dans l'industrie. Pour ne prendre qu'un exemple, mais on pourrait les multiplier dans les domaines de l'automobile, du pétrole ou des systèmes d'armes, les avionneurs utilisent déjà couramment des méthodes de calculs approchés pour améliorer le dessin de l'avion. Le nombre de maquettes testées par Boeing a ainsi été divisé par dix entre deux générations d'avions. Des méthodes approchées, comme la simulation des grandes échelles de la turbulence dans le domaine des moteurs, peuvent, de plus, donner des résultats intéressants avec une machine de 100 teraflops.

La position des grands industriels français utilisateurs de calcul intensif mérite d'être approfondie mais semble cependant dégager d'ores et déjà des caractéristiques intéressantes. Pour des raisons de sécurité et de confidentialité, ils ne sauraient les mutualiser sur de grosses machines ouvertes. Le fait de savoir qu'un industriel procède à des calculs sur un domaine ou un ensemble de domaines est déjà en soi une information qui relève de l'intelligence économique.

Par ailleurs, la brièveté de la durée de vie des machines utilisées pour le calcul scientifique (typiquement moins de cinq ans) induit des coûts de portage récurrents et élevés pour les industriels.

Pour toutes ces raisons, les industriels ne renonceront pas à leurs moyens propres. Ils mettent d'ailleurs en doute la mutualisation des moyens réalisée au CCRT comme principe global, même s'ils utilisent ce dernier à la marge. Ce qu'ils attendent surtout de la sphère publique est un accès facile à un réseau d'experts qui puisse les éclairer dans leurs choix de logiciels et de machines.

Cette vue d'ensemble suggère d'une part de promouvoir la simulation numérique de manière large pour qu'elle puisse être reprise en interne par les entreprises, d'autre part de promouvoir la coopération entre le monde de la recherche académique et celui de la R & D industrielle autour des modèles avancés.

En revanche, il faut éviter le travers de rendre la croissance des centres de calcul intensif destinés à la recherche publique dépendante de la clientèle industrielle. En effet une fois son projet mûr, une entreprise aura une tendance affirmée à internaliser l'exploitation informatique correspondante en son sein, en l'accompagnant éventuellement d'un déploiement au plan mondial calquée sur sa structure globale.

VIII- Externaliser le calcul scientifique ?

Les directeurs des grands centres de calcul français admettent qu'il serait certainement possible d'effectuer un certain nombre de calculs de bas niveau dans le commerce plutôt que de les laisser saturer les grosses machines.

Cela dit, la notion de *computing on demand*, telle que popularisée par IBM, correspond en général à la satisfaction d'un besoin à court terme (3 mois). IBM a indiqué au CINES l'ordre de grandeur de 1 €/heure/processeur. Cela situe le coût du teraflop aux alentours de 800 k€/an. Pour ce prix on dispose d'un ordinateur "nu", seulement doté de son système d'exploitation et exploité par un opérateur, dans une usine IBM. Le prix n'inclut ni capacité de stockage, ni sauvegardes, ni licences logiciel. A cela, il faut ajouter le coût des liaisons nécessaires au raccordement de l'utilisateur au site IBM³⁵. Il en résulte incontestablement une grande complexité de gestion. Au vrai, l'offre d'IBM semble être encore embryonnaire et s'affiner au fur et à mesure des demandes du CINES. En tout état de cause, elle paraît chère³⁶. C'est en fait une offre d'écrêtage de charge basée sur l'utilisation de machines en attente de livraison au client. Il est sans doute un peu tôt pour la considérer sérieusement.

Un point important est toutefois qu'on ne peut pas prétendre sérieusement disposer d'un savoir-faire national si on n'a pas en propre de machine et qu'on se contente d'acheter du temps de calcul à des prestataires extérieurs. Ces équipes de haut niveau, indispensables pour faire un minimum de prospective et prévoir les tendances de l'informatique scientifique, ne peuvent en effet se créer qu'autour de machines.

Quant à l'infogérance, si elle est environ 20 % moins chère qu'un service informatique interne, elle ne marche bien que pour des applications informatiques standard, typiquement des programmes de gestion, mais pas pour des besoins très spécifiques comme ceux du calcul scientifique. Si la question ne se pose donc pas pour les grands centres, elle pourrait en revanche être examinées pour les mésocentres.

IX- Calcul vectoriel, calcul parallèle ou grilles ?

La nature des besoins en calcul intensif diffère selon les applications. *L'intensif* peut venir :

- du nombre important des données à gérer et à prendre en compte pour acquérir une bonne connaissance d'un phénomène et en tirer les conséquences, comme par exemple pour l'analyse des bases de données ; le rapport (quantité d'opérations et d'échange de données au cours du traitement)/(volume de données à traiter) est faible ;
- du nombre important de calculs à effectuer ; ce peut être le cas par exemple pour des calculs de moyennes statistiques, le nombre de données pouvant être très important. Le rapport (quantité d'opérations et d'échange de données au cours du traitement)/(volume de données à traiter) est modéré ;
- du nombre important d'inconnues, toutes interdépendantes, qui entrent dans la modélisation du phénomène, comme en climatologie ; le rapport (quantité d'opérations et d'échange de données au cours du traitement)/(volume de données à traiter) est grand.

³⁵ A titre indicatif, le coût d'une liaison entre le CINES et l'usine IBM de Montpellier (11 km seulement) est d'environ 100 k€/an.

³⁶ 2,5 teraflops tout montés à l'IN2P3 ont coûté 700 k€

Certaines architectures de calculateurs sont à l'évidence mieux adaptées au traitement de certains problèmes que d'autres. C'est notamment le cas des architectures dites *vectérielles*, très en vogue dans la dernière décennie et bien adaptées à certains programmes de climatologie.

Ces architectures sont apparues dans les années 1970 avec les premiers ordinateurs Cray. Elles consistent à manipuler des vecteurs de nombres en virgule flottante qui peuvent être chargés directement de la mémoire dans des registres vectoriels pour être traités ensuite par l'unité arithmétique du calculateur, au lieu de manipuler des nombres individuels, comme c'est le cas dans les calculateurs dits *scalaires*.

Une seule machine vectorielle figure aujourd'hui dans les dix premières du TOP 500, l'Earth Simulator japonais, et seule l'industrie japonaise, NEC notamment, pousse encore vraiment ce type d'architecture, très consommatrice d'électricité, de surfaces et de climatisation et très pointue en termes d'applications qu'elle peut traiter. En revanche, on note l'émergence dans le TOP 500 des machines massivement parallèles basées sur des architectures scalaires. La mode du calcul vectoriel, tel que l'ont illustré les machines Cray ou NEC, semble donc aujourd'hui passée. L'évolution future des architectures de grands calculateurs sera certainement déterminée par leur capacité à rentabiliser sur le marché les investissements considérables que requièrent leur développement. Même s'il est encore risqué d'annoncer sa disparition définitive, la généralisation du parallélisme paraît une tendance lourde de l'informatique, à laquelle il faut se préparer.

Cette évolution concerne essentiellement la communauté des climatologues dont les travaux, en France, reposent largement sur l'utilisation du NEC SX-5 de l'IDRIS. Pour autant, s'il n'y a pas aujourd'hui d'autre choix que de remplacer à brève échéance cette machine vieillissante, il convient aussi de se préparer dès maintenant à l'échéance suivante, vers 2010, où pourraient bien ne plus être disponibles que des machines parallèles à base de processeurs scalaires.

Le concept de *Grid Computing* ou de *grilles* se présente de façon totalement différente. Son heure de gloire est arrivée parce que la technologie a juste atteint le niveau de maturité où elle est plus ou moins utilisable dans certains contextes d'applications spécialisées. On sait que 40 % à 80 % de la puissance de traitement des ordinateurs en service aujourd'hui reste inutilisée. Construites autour de réseaux de télécommunications, les grilles devraient justement permettre une meilleure utilisation de cette puissance disponible. Mais la mise en œuvre de cette distribution de la capacité va de pair avec la distribution des traitements et du stockage et la circulation des données. L'idée est simple, voire séduisante, mais sa mise en œuvre concrète réclame des avancées spectaculaires dans de nombreux domaines car pour être acceptée les grilles devront se montrer aussi fiable qu'un grand ordinateur individuel. C'est l'objectif du programme GRID 5000 que l'INRIA développe dans le cadre d'une action concertée incitative (ACI). Pour autant, une grille distribuée restera handicapée par la réalité physique pour ces types d'applications. En effet le temps de latence, c'est à dire celui qui sépare l'envoi d'une donnée de sa réception, est un facteur déterminant pour les performances d'un grand code scientifique unique. Des constructeurs de machines parallèles, comme SGI, se fixent ainsi pour objectif de le ramener à moins d'une microseconde entre deux processeurs d'ici 5 ans. Or en une microseconde, un signal électrique ne parcourt qu'environ 300 mètres, distance infime à l'échelle d'un réseau d'ordinateurs à l'échelle nationale, a fortiori paneuropéenne ou mondiale.

On peut toutefois penser que les *grilles* auront à terme un rôle structurant pour la recherche au niveau international, pour le calcul, les bases de données, le travail collaboratif et certaines activités expérimentales. Il convient donc de mettre en place et d'afficher une politique dans ce domaine, qui permette la satisfaction des besoins de la communauté scientifique par des projets pilotes et tire profit des résultats de la recherche dans les STIC. La réflexion stratégique doit également permettre de se préparer aux effets structurants de la mise en place de grilles sur l'organisation de la recherche, ce qui impliquera la constitution de communautés virtuelles internationales, la mise en réseau des connaissances et expertises, et la possibilité de délocaliser certains équipements mutualisables.

Le concept de grille paraît plus intéressant dans le monde de la gestion que dans celui de la performance. Il n'est pas adapté à la production lourde. Il n'y a d'ailleurs pas de grille dans le TOP 500. En fait, les projets de grilles apparaissent plus comme des projets de partage de données (décodage du génome, ...) que de vrais projets de calcul. Elles ne résolvent pas tout et en tout cas ne remplaceront jamais un grand centre de calcul. Cela dit, en tant que nouvelle technologie, elles méritent que la France s'y intéresse et il convient de soutenir les initiatives européennes et internationales du CC-IN2P3 dans ce domaine, notamment dans le cadre du programme EGEE.

X- Eléments d'une politique

Plusieurs arguments militent pour que, sans pratiquer systématiquement la course à la puissance, la France reste bien classée dans la communauté mondiale du calcul scientifique :

- la concurrence internationale est particulièrement vive dans le domaine de la recherche.
- derrière cette concurrence se cachent des enjeux de pouvoir³⁷ : des négociations internationales de grande envergure comme celle du protocole de Kyoto s'appuient sur les résultats de modélisations que les scientifiques mettent à la disposition des politiques ; peut-on imaginer un seul instant que les négociateurs européens s'en remettent entièrement à des données fournies par des experts américains ou japonais ? Il est donc essentiel, sinon pour la France du moins pour l'Europe, de garder une capacité autonome d'expertise, voire de contre-expertise, dans les domaines nécessitant de grands calculs (climatologie, nucléaire civil, biologie, nanotechnologies, ...).
- les progrès en théorie et en description des systèmes complexes suivent dans une large mesure ceux des ordinateurs ; ne pas mettre à disposition des chercheurs les outils le plus en pointe revient à les handicaper durablement dans la compétition internationale.
- le maintien d'équipes de classe internationale dans le domaine du calcul scientifique, aussi bien en matériel qu'en logiciel, n'est possible que si celles-ci peuvent se regrouper autour de machines qui figurent dans les dix ou vingt premières du TOP 500.
- enfin, la puissance de calcul reste toujours à la base du poids de chaque partenaire dans un accord international ; si donc on ne veut pas laisser marginaliser les centres de calcul français face à leurs homologues européens, il faut les équiper en conséquence.

Si la France veut compter parmi les principaux foyers mondiaux de recherche, elle doit impérativement, dans le domaine du calcul intensif, réformer son système pour mieux coordonner ses efforts au niveau national, développer la prospective et y consacrer des moyens financiers réguliers, étant précisé que les considérations de politique industrielle ne rentrent pas dans le cadre de cette mission.

Les rapporteurs estiment donc qu'il convient de lancer rapidement en France une action destinée à :

- mettre en place une coordination stratégique au plan national ;
- programmer le rattrapage du retard national vis-à-vis du Royaume Uni et de l'Allemagne au niveau des moyens de calcul intensif ;

³⁷ Même si les applications militaires se situaient en dehors du champ de la mission, il est intéressant de citer ce propos de M. D. Verwaerde (CEA/DAM) qui a bien exprimé la réalité des synergies entre les domaines civil et militaire dans son exposé au forum de l'ORAP en juin 2004 : *La pérennité de la simulation repose sur une stratégie en 3 points : 1) Diffuser le plus largement possible le recours à la simulation dans la recherche et l'industrie, 2) Développer avec l'Université, les formations et la recherche dans le domaine de la simulation, 3) Maîtriser toutes les briques technologiques du calcul à hautes performances via des collaborations avec des laboratoires universitaires et des industriels de l'informatique.*

- participer activement aux initiatives européennes en matière de calcul intensif ;
- développer les compétences nationales et à renforcer significativement les coopérations scientifiques interdisciplinaires, ainsi qu'entre recherche et industrie ;
- garantir un niveau de financement convenable et pérenne du calcul scientifique.

Il est clair que la nature des enjeux qui ressortent de cette liste, dépassent largement la stratégie d'équipement d'un centre isolé.

Recommandation 1 : mettre en place un comité stratégique du calcul scientifique

Il ressort des auditions effectuées par la mission un besoin ressenti par l'ensemble des acteurs de l'affirmation d'une véritable politique nationale du calcul scientifique qui donnerait un éclairage à moyen terme à la nécessaire évolution des capacités de calcul nationales. Toutes les personnes rencontrées mettent d'ailleurs en avant l'existence dans la plupart des grands pays d'une organisation, comité ou conseil scientifique, qui produit des analyses et des avis concernant le calcul scientifique, notamment au Royaume-Uni avec le *Council for Science and Technology*³⁸. Il semble que l'absence de ce type de structure explique largement le caractère chaotique de l'évolution de nos capacités de calcul. La proposition de mettre en place un processus de pilotage stratégique permet de prendre en compte le contexte en évolution rapide, la multiplicité des organismes concernés au plan national et le besoin de définir des plans pluriannuels. Le réalisme budgétaire indique qu'il faudra dans ce cadre prendre des décisions difficiles, et veiller au suivi des programmes.

Il est donc recommandé que le ministre délégué à la Recherche nomme un **comité stratégique du calcul scientifique**. La composition de ce comité stratégique est un point crucial en raison des conflits potentiels d'intérêt auxquels peuvent se trouver confrontés ses membres. Elle devra donc être soigneusement étudiée. Ce comité pourrait compter une quinzaine de personnes, dont trois étrangers et deux industriels utilisateurs de calcul intensif³⁹, nommées *intuitu personae* par le ministre sur des listes trois fois plus larges proposées par les grands organismes de recherche. Ce comité pourrait être renouvelé tous les trois ans.

Le comité stratégique du calcul scientifique pourrait aussi devenir un sous-comité du Haut conseil pour la recherche et l'innovation (HCRI). Il pourrait aussi conseiller l'Agence nationale de la recherche (ANR), avec laquelle il devra en toute hypothèse se coordonner étroitement. Il pourrait, enfin, en tant que de besoin, conseiller l'Agence pour l'innovation industrielle (AII).

Le secrétariat de ce comité serait assuré par la direction de la Recherche qui serait l'exécutif des décisions du ministre.

Ce comité aurait pour mission :

- de remettre tous les ans au ministre délégué à la Recherche un rapport de synthèse, scientifique et budgétaire, de ses travaux à une date permettant son intégration dans la nouvelle procédure budgétaire ; il devrait donc, dès son installation, définir les indicateurs de performance du programme national de calcul intensif et les diffuser auprès de l'ensemble des communautés scientifiques concernées.

³⁸ <http://www.cst.gov.uk/>.

³⁹ Il est difficile d'y intégrer des industriels constructeurs de moyens de calcul sans risquer d'introduire des distorsions de concurrence.

- d'instituer par discipline et par établissement⁴⁰ des **groupes permanents d'études prospectives** visant à faire un état des besoins actuels en calcul intensif et de leur évolution ; certaines disciplines, comme la biologie, répartie sur un grand nombre d'organismes, ou les nanosciences, de nature pluridisciplinaire, pourraient se voir doter d'une organisation particulière pour tenir compte de cette diversité et des spécificités des sous-disciplines face au développement des techniques de simulation⁴¹ ; le comité stratégique du calcul scientifique serait destinataire des études prospectives et chargé d'en établir une synthèse ;
- de transformer les orientations issues de la prospective en schéma directeur et en plan à moyen terme de renouvellement des équipements ;
- de conseiller l'administration de la recherche et de l'enseignement supérieur sur les mesures d'accompagnement nécessaires pour optimiser l'utilisation du calcul intensif par les disciplines (formation des doctorants des disciplines utilisatrices, modernisation des codes et des algorithmes, techniques de modélisation...) ;
- de donner un avis sur les cahiers des charges et le choix des fournisseurs dans les appels d'offres des centres de calcul intensif à vocation généraliste ;
- d'introduire en France une approche projet dans le domaine du calcul scientifique, comme cela se pratique aux Etats-Unis.

Recommandation 2 : combler le retard français en calcul intensif

La qualité de la recherche française exige une puissance de calcul intensif comparable à celle de ses homologues européennes. Il est indispensable et urgent que l'Etat prenne des décisions sur le financement nécessaire et le planifie dans la durée, de façon qu'il soit assuré, régulier et stable. Des opérations ponctuelles sont insuffisantes. La régularité et la stabilité sont d'ailleurs peut-être plus importantes que la valeur absolue des sommes affectées au calcul intensif.

Un objectif raisonnable est de rattraper nos voisins allemand et britannique qui d'après la loi de Moore, devraient disposer d'une puissance de calcul de l'ordre du petaflop vers 2011, alors qu'actuellement la nôtre n'est que la moitié de la leur⁴². A cet égard, il est urgent de laisser le CINES lancer son appel d'offres en suspens depuis deux ans et l'IDRIS renouveler son calculateur vectoriel, sous les réserves indiquées pp. 22 et 29. L'échéancier pourrait être le suivant :

⁴⁰ Une lettre du ministre aux instances dirigeantes des organismes de recherche montrerait l'intérêt que porte l'Etat à ce domaine.

⁴¹ Un tel comité existe déjà de manière permanente au CEA, ce qui permet de nourrir les plans glissants de l'organisme. Il faudrait l'instituer ou le relancer, et le rendre permanent, au CNRS, à l'INSERM, à l'INRA et dans tout autre organisme de recherche concerné par le calcul intensif.

⁴² Fin 2004 les calculateurs "academic" + "research" français comptaient pour 13 teraflops dans le TOP 500, les allemands pour 30 et les britanniques, hors ECMWF (voir p. 15) pour 25. La loi de Moore appliquée sur la base d'un doublement tous les 21 mois conduit alors exactement à 412 teraflops pour la France, 920 teraflops pour l'Allemagne et 785 teraflops pour le Royaume-Uni, toujours hors ECMWF, en 2010.

année	puissance (teraflops)
2004	13
2005	40
2006	80
2007	150
2008	250
2009	400
2010	650
2011	1000

Rattrapage du retard français en matière de puissance de calcul

L'effort financier pour atteindre cet objectif demeure modéré. La simple application de la loi de Moore au CINES et à l'IDRIS, *minimum minimorum* pour leur éviter de perdre toute visibilité sur la scène internationale, exige qu'ils disposent chacun d'un budget annuel régulier d'investissement de 5 M€ L'effort de rattrapage indiqué ci-dessus conduirait *grosso modo* à doubler cet effort et à passer de $2*5 = 10$ M€ à $2*5 + 10 = 20$ M€ par an.

La mission suggère de placer cet effort dans un cadre de TGE (Très Grand Equipement), ce qui facilitera la planification de son utilisation scientifique et permettra d'appuyer les réalisations sur les compétences d'exploitation et d'utilisation.

Recommandation 3 : la structuration des acteurs du calcul intensif

La mission ne trouve pas déraisonnable que la France dispose d'au moins deux grands centres de calcul généralistes physiquement distincts. La répartition en deux sites peut aussi se justifier pour éviter une paralysie en cas d'incidents majeurs mais surtout pour offrir des services complémentaires. Il semble en effet clair que pour répondre aux demandes variées des chercheurs en calcul scientifique plusieurs types machines doivent être proposées. Dans un paysage industriel et technologique en évolution permanente, il n'est pas raisonnable non plus de ne dépendre que d'un seul fournisseur. Les rapporteurs recommandent donc de ne pas modifier la répartition géographique actuelle des grands centres de calcul français.

En revanche, ils ne verraient que des avantages à les réunir dans une structure juridique unique⁴³ à commandement unifié. S'il n'est pas question de modifier l'organisation physique des différents centres et leur fonctionnement, il s'agit en revanche de produire une politique commune de développement et de fournir un interlocuteur unique aux pouvoirs publics, nationaux ou européens et de garantir ainsi la pérennité et l'excellence des moyens de calcul nationaux. Il est entendu que la vocation du CC-IN2P3 au service d'une communauté particulière, celle de la physique des particules, le laisse a priori à l'écart de cette structuration.

Il est important que la solution proposée concilie le besoin d'harmonisation nationale de la politique du calcul scientifique et l'efficacité de la gestion au quotidien des centres de calcul.

⁴³ Rappelons que, pour l'instant, le CCRT du CEA ne dispose pas de la personnalité juridique, pas plus que l'IDRIS, simple unité propre de service du CNRS. Seul le CINES dispose de cette personnalité, du fait de son statut d'EPA.

Dans cette direction, la mission propose, en ligne avec les conclusions du rapport de l'Inspection générale des Finances et de l'Inspection générale de l'administration de l'Education nationale et de la Recherche sur les grands équipements, de regrouper tout ou partie des acteurs du calcul scientifique dans une structure juridique souple, telle qu'**une société civile du calcul intensif**, dont seraient actionnaires les établissements de recherche (organismes et universités) et les ministères concernés⁴⁴. Ce schéma peut se concevoir selon deux variantes qui ne sont au demeurant pas neutres au regard de leurs implications exposées plus loin sur le projet européen (voir p. 29) :

- une version extensive qui regroupe l'ensemble des centres existants (CINES, IDRIS et CCRT) dans une structure juridique unique ;
- une version limitée qui tient compte des spécificités du CCRT : un fonctionnement coordonné avec la partie militaire du CEA, dont la compétence en matière de gestion de grosses machines de calcul est importante et la présence dans cette *coopérative* informelle de plusieurs industriels ; la structuration de l'ensemble français pourrait alors se faire par la création de deux sociétés civiles, l'une constituée autour du CINES et de l'IDRIS, l'autre autour du CCRT, ayant des participations croisées l'une dans l'autre, mais chacune gardant ses relations d'origine.

Il serait sans doute utile d'intégrer aussi les principaux mésocentres dans le dispositif de programmation générale en raison de leurs rôles de centres de mise au point des programmes qui s'exécutent ensuite sur les grands centres et aussi de nœuds de grilles, existants ou potentiels. La mission n'a pas eu le temps d'explorer en détail cette possibilité qui devrait faire l'objet d'un approfondissement par la direction de la recherche.

Recommandation 4 : renforcer la coopération européenne

Renforcer la coopération européenne, par la création et la gestion en commun d'un très grand calculateur est une opération nécessaire pour la recherche européenne.

Il faut cependant bien comprendre que les machines installées dans le cadre de la coopération européenne sont essentiellement destinées à des programmes qui ne pourraient pas être exécutés sur les capacités nationales même fortement accrues.

La demande en calculs (de toutes dimensions) va connaître une expansion très importante sans qu'il soit possible de la chiffrer de façon même approximative. Il faudra donc également faire croître les capacités nationales et mieux répartir l'exécution des programmes entre les divers niveaux de machines.

Le rapport annexe présente la prospective du calcul intensif pour la recherche française. Il en ressort clairement que la machine européenne ne sera sollicitée que par certains projets de très grande envergure. Mais même si la recherche française ne représente qu'une fraction de la recherche européenne, il est probable que la machine, a-priori au service de l'ensemble des pays de l'Union européenne, frôlera rapidement la saturation, sous l'effet des demandes d'exécution de grands challenges. Il est donc certain que les autres grands défis scientifiques à relever par la recherche française ne pourront pas tous s'exécuter sur la machine européenne. Ils ont donc besoin de capacités de calcul fortement accrues par rapport à la situation actuelle. Il n'y a donc pas substituabilité entre la machine européenne et les capacités nationales.

⁴⁴ La société civile, propriétaire des machines, pourrait les mettre à disposition des centres existants après un appel d'offre et la signature d'une convention d'utilisation conforme à un cahier des charges défini par la société.

Une mise en place rapide de l'organisation recommandée ci-dessus permettrait de gérer au mieux des intérêts nationaux, en particulier à l'occasion des négociations dans la phase actuelle de préparation du 7^{ème} PCRD. Il faut également soutenir les actions de coopération internationale de manière à maintenir la compétitivité des recherches françaises dans les domaines qui reposent sur le calcul, et à bien nous positionner dans les projets que mettent en place nos partenaires.

La mission recommande également de poursuivre les négociations tripartites entre l'Allemagne, la France et le Royaume-Uni sur la mise en place d'une infrastructure européenne de calcul de très grande puissance, financée pour partie par la Commission européenne dans le cadre du programme ERA-Net, avec la volonté d'aboutir. C'est la seule voie qui permette à nos chercheurs de disposer à terme et de façon durable de moyens de calcul comparables à ceux de leurs homologues américains ou japonais.

Ce pourrait alors être une mission du comité stratégique de définir une doctrine d'emploi des machines nationales d'un côté, de la machine européenne de l'autre, de façon à optimiser l'utilisations des unes et de l'autre. La ligne de partage pourrait être que seuls les très grands codes, une fois mis au point, auraient accès à la machine européenne. Leur mise au point et l'exécution de programmes moins gourmands en ressources se feraient, elles, sur les machines nationales.

Sur la base des indications données par la direction de la Recherche, les dépenses d'investissement monteraient à 200 M€ tous les 2 ans à partir de 2008 et seraient supportées pour 35 % par la Commission européenne, à 15 % par des industriels ou d'autres pays de l'Union européenne, à 25 % par le pays hôte et à 5 % par chacun des deux autres partenaires et trois autres pays associés. Les frais de fonctionnement sont évalués globalement à 20 M€ par an les deux premières années, 40 M€ par an les deux années suivantes et 60 M€ les deux années d'après, supportés à égalité entre tous les utilisateurs, soit a priori 1/6 pour la France. Cela dit, il est certainement possible pour la France de mutualiser au moins en partie ces coûts avec ceux de l'IDRIS ou du CEA, par exemple, notamment pour ce qui est du personnel d'exploitation. Si on se limite à l'aspect investissement, l'échéancier envisagé dans le cadre des négociations en cours entre les trois pays est donc le suivant :

(M€)	2008	2009	2010	2011	2012	2013	Total	Moyenne
Investissement	200		200		200		600	100
Commission (35 %)	70		70		70		210	35
Industriels (15 %)	30		30		30		90	15
Pays A⁴⁵	50		10		10		70	11
Pays B	10		50		10		70	11
Pays C	10		10		50		70	11
Associés D, E, F	30		30		30		90	15
Total	200		200		200		600	100

Si la France veut donc jouer cette carte, c'est 17 ou 18 M€ dont 12 M€ d'investissements, qu'elle doit consacrer annuellement à ce projet à partir de 2008.

Dans sa configuration actuelle, le projet tripartite a notamment pour ambition de satisfaire les besoins de la communauté climatologique et donc de rendre sans objet le projet ENES d'ordinateur européen dédié à la climatologie, sur le modèle de l'Earth Simulator. Cela suppose toutefois que soient adoptées rapidement les mesures recommandées par ailleurs en matière de parallélisation des grands codes de climatologie, tant il est vraisemblable que, si le projet tripartite voit le jour, il recourra à des machines parallèles.

⁴⁵ A, B et C désignent globalement l'Allemagne, la France et le Royaume-Uni, aucune discussion n'ayant encore eu lieu pour déterminer l'ordre dans lequel les calculateurs seraient installés dans chacun de ces trois pays. D, E et F désignent les autres pays utilisateurs de ces moyens de calcul.

Pour ce qui est des structures d'accueil au plan français, les deux possibilités évoquées à la recommandation 3 (voir p. 27) présentent des différences quant à l'articulation avec le projet européen. La version extensive accueillerait facilement la partie française du projet européen. Il appartiendrait alors au conseil d'administration de la société de définir les modalités précises de son intégration dans l'ensemble national. Comme la structure juridique européenne serait vraisemblablement de droit privé, l'articulation se ferait probablement par prise de participation de la société civile française dans la société européenne. Dans la version limitée, l'Etat devrait procéder lui-même aux arbitrages pour décider lequel des deux centres français (structure commune CINES-IDRIS ou CCRT) accueillerait la machine européenne, ce qui au demeurant peut se faire par appel d'offre. Une seule des deux sociétés prendrait alors une participation dans l'ensemble européen, ce qui risque de marginaliser l'autre.

Enfin, en ce qui concerne les grilles, la mission suggère de confier au comité stratégique le soin de suivre les progrès de la recherche et le développement des technologies, en particulier en évaluant les projets DEISA et EGEE, et de proposer des déploiements pilotes, correspondant à des communautés ciblées dans le monde de l'enseignement supérieur et de la recherche.

Recommandation 5 : faire face à la possible disparition du calcul vectoriel

Il faut aider la communauté des climatologues à se préparer à l'éventualité de la disparition de machines vectorielles, en commençant par une analyse scientifique de la question. Cette mission pourrait être confiée à des experts reconnus (CERFACS, CEA-DAM, EDF études et recherches, ORAP, EADS) chargés de faire l'état des lieux sur le passage du vectoriel au parallèle, d'en évaluer le coût, la durée et de proposer des sources de financement. Elle devrait également regarder la possibilité de renforcer temporairement les équipes scientifiques concernées par du personnel en sous-traitance pour mener à bien la parallélisation des codes vectoriels existants, de façon à soulager les chercheurs et leur permettre de poursuivre leurs travaux. Il est d'autant plus urgent d'agir que l'effet d'une telle mesure ne se fera sentir qu'à moyen terme, typiquement 3 ou 4 ans plus tard, dont peu avant l'échéance de 2010.

Si la décision est prise d'abandonner la voie des calculateurs vectoriels, il faudra alors annoncer clairement et rapidement la couleur : l'IDRIS sera autorisé à remplacer en 2005 sa machine vectorielle NEC SX-5 qui date de 2000, de façon à laisser le temps nécessaire à l'adaptation des grands codes vectoriels entre maintenant et 2009 ou 2010, mais ce sera la dernière machine vectorielle à usage scientifique achetée en France.

Cette réflexion stratégique et technique sur l'évolution des architectures de calculateurs mériterait d'ailleurs d'être pérennisée tant il est difficile de l'appréhender longtemps à l'avance, en raison de l'imbrication forte entre considérations techniques et commerciales.

Recommandation 6 : développer les synergies en matière de logiciels

Le calcul scientifique est une activité qui requiert des machines mais aussi des algorithmes et des logiciels. Sur ces deux aspects, malgré la qualité exceptionnelle de ses équipes, la communauté française souffre. La situation exige un effort urgent.

Il faut pousser les feux sur les technologies algorithmiques. Le CERFACS pourrait servir de cristal de base d'autant qu'il possède une structure de société civile à laquelle il serait facile que les organismes de recherche ou même le ministère puissent adhérer. Faut-il en créer une filiale (pour profiter de son savoir-faire unanimement reconnu) en région parisienne et l'inclure dans un pôle de compétitivité à créer en Ile-de-France Sud ? La question est à examiner. En tout cas, le CERFACS pourrait participer à une organisation du calcul scientifique méditerranéen, de Barcelone à Cineca (Bologne) en Italie en passant par le CINES, et ITER qui aura des besoins en calcul très importants, s'il est implanté à Cadarache.

Plus précisément, il faut :

- améliorer la méthodologie de développement des grands codes, notamment par le développement de bibliothèques de logiciels et d'ateliers de génie logiciel et le développement de logiciels de base, notamment en capitalisant sur l'expérience du CEA en matière de logiciels libres ; il s'agit de consolider les résultats de la recherche dans des codes mutualisés fiables, optimisés et éventuellement susceptibles de valorisation industrielle ; la bonne conduite de ces opérations passe par la coopération avec des chercheurs et ingénieurs de recherche en mathématiques appliquées et informatique ; par la sélection rigoureuse des projets à mener vers ce niveau ;
- favoriser le développement de nouvelles applications dans des domaines tels que ceux des nanotechnologies et des biotechnologies et, plus généralement, dans tous les nouveaux domaines d'utilisation du calcul scientifique ;
- expérimenter et valider, dans des conditions d'utilisation réalistes, des architectures de calculateur innovantes et compétitives ; les logiciels concernés pourront comprendre une infrastructure de grille et de communications rapides, la gestion et la manipulation de données, des logiciels de base de compilation avancée et de gestion de données distribuées.

La mission estime notamment qu'il est possible de tirer un meilleur parti de l'exceptionnel regroupement de compétences en Ile-de-France Sud autour de l'université de Versailles-Saint-Quentin-en-Yvelines, de l'INRIA, du plateau de Saclay (CEA, Ecole Polytechnique, Ecole Supérieure d'Electricité, Ecole Supérieure d'Optique, université Paris XI Orsay), de l'IDRIS, et des activités civiles du CEA à Bruyères-le-Châtel (CCRT, TERATEC, ainsi que des industriels qui ont déjà des habitudes de travail avec le CEA : HP, Bull, CS, SNECMA, EDF, auxquels pourraient s'adjoindre l'ONERA, Simulog et Dassault Aviation. Les rapporteurs voient également bien un axe Lyon-Grenoble autour de l'IN2P3 et du CERN.

Recommandation 7 : développer un réseau d'experts

Si le calcul scientifique est une activité qui requiert des machines et des logiciels, il exige aussi des hommes. Les grands projets de calcul intensif requièrent la mise en œuvre de compétences pluridisciplinaires. En dehors du domaine d'application, il faut maîtriser certains aspects de génie logiciel, d'algorithmique et d'analyse numérique. Il faudra poursuivre les actions entreprises, en particulier pour la formation des chercheurs, et les compléter dans les domaines de la réalisation de grands codes et du génie logiciel.

La France doit disposer d'experts dans le domaine du calcul intensif pour répondre à trois types de besoins :

- un besoin de veille technologique et de prospective sérieuses, voire d'intelligence économique, dans un domaine industriel et technologique en mutation extrêmement rapide ;

- un besoin de disposer d'experts en architecture de systèmes informatiques, dans la mesure où la notion de supercalculateur universel qu'on a connue (Cray, Fujitsu, ...) a quasiment disparu au profit de solutions bâties plus ou moins au coup par coup à base de briques du commerce (processeurs, réseaux d'interconnexion, unités de stockage) en fonction des besoins exprimés par l'utilisateur ;
- un besoin de disposer d'experts en programmation capables de développer les outils informatiques nécessaires à une exploitation efficace des machines et de les mettre à la disposition de la communauté des utilisateurs qui n'est plus en mesure de développer seule ses grands codes (ateliers logiciels) ; ces développements se font souvent en coopération internationale sur la base de logiciels libres.

Le volet de la coopération entre recherche fondamentale et recherche d'intérêt technologique devra également être développé. Il convient de s'appuyer sur l'expérience du CEA, notamment autour du CCRT.

Cette politique doit donc également faire émerger un réseau d'experts destiné à accroître la réactivité du système, à rendre efficace la veille technologique et l'intelligence économique du dispositif. Ceci demandera un effort supplémentaire d'investissement constant au cours du temps, en raison de l'évolution rapide de la technologie, qui fait que les matériels se périment au bout de 4 ans environ.

Ce réseau permettrait de mieux valoriser les investissements individuels dans le logiciel libre. Pour cela, le pilotage des communautés de logiciel libre sélectionnées pourrait se faire en concentrant sur elles les contributions techniques du meilleur niveau. Cela suppose évidemment de valoriser l'activité des chercheurs impliqués. Au demeurant, des politiques de ce type sont déjà mises en œuvre aux Etats-Unis, par exemple dans le cadre du plan OSSODA⁴⁶, ou par des firmes privées comme IBM⁴⁷. En France, le CEA et Bull⁴⁸ recourent aussi à une telle approche, et ont déjà obtenu des succès sensibles dans des projets internationaux.

Plus généralement, il est essentiel que le système de l'Education nationale favorise la formation au calcul intensif, y compris dans les disciplines utilisatrices. Il est fondamental en effet que les jeunes diplômés et les jeunes chercheurs apprennent à voir grand en matière d'informatique et ne raisonnent pas uniquement en termes d'ordinateur personnel comme solution à leurs besoins de calcul.

Recommandation 8 : accroître et pérenniser les moyens financiers du calcul intensif

Le tableau ci-dessous résume les investissements nécessaires dans les grands centres de calcul généralistes. Il n'inclut ni le CCRT, faute de données suffisamment précises de la part du CEA, ni les mésocentres, pour les raisons indiquées plus. Il ne tient pas compte non plus du CC-IN2P3, dont le caractère généraliste ne paraît pas suffisamment affirmé. En revanche il inclut les mesures d'accompagnement. Ces dernières comprennent notamment le financement de la parallélisation des grands codes de climatologie, indispensable pour faire face à la possible disparition des calculateurs vectoriels à l'horizon 2010. Le tableau tient compte également une estimation du projet tripartite sur la base de ce qu'on en sait aujourd'hui. Il est toutefois probable que certaines économies d'échelle sont envisageables, par exemple en hébergeant le calculateur français au CEA ou à l'IDRIS. On peut ainsi bénéficier d'infrastructures (bâtiments, climatisation) existantes, voire de personnel d'exploitation.

⁴⁶ *Open Source Software Development Acceleration.*

⁴⁷ <http://www-128.ibm.com/developerworks/opensource/library/os-welcome.html>.

⁴⁸ <http://www.clusterfs.com/pr/2004-11-11.html> et <http://www.clusterfs.com/partners.html>.

Opération (M€)	2005	2006	2007	2008	2009	2010
CINES	5	5	5	5	5	5
IDRIS	5	5	5	5	5	5
rattrapage	10	10	10	10	10	10
Accompagnement	4	4	4	4	4	4
Projet D-F-UK ⁴⁹				18	18	18
Total (M€)	24	24	24	42	42	42

Besoins en investissements 2005-2010

Il convient de voir que, d'une part, ce tableau ne constitue pas une demande entièrement nouvelle et que, d'autre part, il n'est pas entièrement à charge du budget du ministère délégué à la Recherche. Sur le premier point, il faudra de toute façon bien assurer les investissements indispensables dans les centres de calcul et dans les réseaux (RENATER). D'autre part, le financement de l'IDRIS et celui du CC-IN2P3 sont en principe assurés entièrement sur le budget du CNRS.

Ce que le tableau veut souligner, ce sont les investissements à consacrer au calcul intensif⁵⁰, si la France prétend jouer dans la cour des grands et mettre à la disposition de ses scientifiques des moyens de calcul comparables à ceux de leurs homologues des autres grands pays. **En ordre de grandeur, il n'y aura pas de politique française ambitieuse en matière de calcul intensif si on n'y investit pas 25 M€ par an, cette somme montant à 40-45 M€ par an, quand le projet de grand calculateur européen, auquel la France se doit de participer, aura vu le jour.**

⁴⁹ Y compris les coûts de fonctionnement.

⁵⁰ Il est clair que si la solution proposée dans la recommandation 3 est adoptée, il faudra globaliser l'ensemble de ces sommes dans le budget de la société civile de calcul intensif et ajouter (dans l'hypothèse d'une solution extensive) la quote-part du CCRT.

Annexes

Annexe 1 : Lettre de mission

Annexe 2 : Liste des personnes rencontrées



*Ministère de l'Éducation nationale,
de l'Enseignement supérieur et de la Recherche*

Le Ministre délégué à la Recherche

Réf. : Cab/CHL/VP/n° 49.04

Paris, le 21 SEP. 2004

Monsieur l'Inspecteur Général,

Notre politique nationale dans le domaine du calcul scientifique constitue un élément important témoignant de la capacité de la France à produire une recherche moderne, innovante et de très haut niveau. A ce stade, plusieurs rapports rédigés pour le compte du Ministère de la Recherche ont fait le constat de l'importance de mener dans les prochaines années une action soutenue en faveur du calcul scientifique.

Les acteurs du calcul scientifique français, qui développent actuellement sur financement public des moyens de calculs au service de la communauté scientifique, apparaissent nombreux. Je citerai de manière non exhaustive le CNRS, Météo-France, le CEA, le CINES, l'INRIA, ou encore les universités au travers des « meso-centres ».

Dans l'optique de mener une action forte en faveur du calcul scientifique, il devient donc plus que jamais nécessaire de réfléchir à la coordination à mettre en place entre ces différents centres, ainsi qu'aux avantages ou inconvénients qui pourraient résulter d'une mise en commun accrue de nos moyens de calculs.

Par ailleurs, les coopérations public-privé à l'initiative d'institutions publiques, ou les initiatives privées au services des communautés scientifiques sont des approches novatrices, que je souhaite encourager dans le cadre de la mise en œuvre de la politique française en matière de calcul scientifique. Nous disposons sur ce point d'exemples intéressants, comme le projet TERATEC du CEA, les services « à la demande » proposés par les constructeurs informatiques, ou encore le CERFACS qui associe Météo-France, le CNES, EADS et la SNECMA.

Monsieur Thierry BOSSARD
Chef du Service de l'Inspection
Générale de l'Administration de
l'Éducation nationale et de la Recherche
107 rue de Grenelle
75007 PARIS

.../...

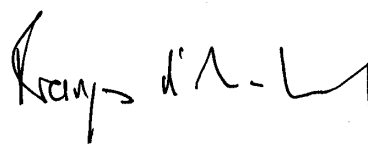
Afin de définir et d'organiser concrètement la politique française dans ce domaine et rationaliser nos choix budgétaires, un groupe de travail sera mis en place, associant les services du ministère chargé de la recherche (Direction de la Recherche et Direction de la Technologie) et les organismes concernés. Je souhaite que le pilotage de ce groupe soit assuré par un membre de l'Inspection générale de l'administration de l'éducation nationale et de la recherche et un membre du Conseil général des technologies de l'information.

Les travaux de ce groupe de travail devront permettre de définir et de préfigurer les structures de concertation et de décision nécessaires à la mise en œuvre de notre politique. Elle devra aborder cette problématique sur le plan de l'identification des besoins, la définition des moyens, leur répartition géographique, et leurs modalités de financement. Le cadre et les attendus du travail du groupe porteront notamment sur les éléments fournis en annexe. Je souhaite être informé mensuellement, et chaque fois que nécessaire, des propositions et des préconisations du groupe afin de pouvoir décider de leur mise en œuvre.

Alors que le gouvernement travaille à la définition de pôles d'excellences, je souhaite que vous portiez une attention particulière à la possibilité de faire émerger en France un pôle de visibilité mondiale dans le domaine du calcul scientifique et de la recherche dans cette discipline.

Un courrier similaire est adressé à Monsieur le Vice-Président du Conseil général des technologies de l'information (CGTI).

Je vous prie de bien vouloir agréer, Monsieur l'Inspecteur Général, l'expression de mes sentiments les meilleurs.

A handwritten signature in black ink, appearing to read 'François d'Aubert', written in a cursive style. The signature is positioned above a horizontal line.

François d'AUBERT

ANNEXE

Le groupe de travail que vous coordonnez établira d'ici la fin du mois de novembre :

- Un inventaire technique et financier (en coûts complets) des moyens existants, avec une visibilité pluriannuelle sur les financements identifiés et les communautés utilisatrices ;
- Une analyse des besoins actuels et futurs, classés par thématique scientifique ;
- Une préconisation de stratégie de développement des moyens, sur le plan géographique (concentration des moyens) et financier (partenariat public-privé pouvant aller d'un co-financement jusqu'à l'achat de services, financement public, modalités de facturation aux équipes scientifiques). Cette préconisation sera argumentée par une analyse comparative des stratégies adoptées par les principaux acteurs mondiaux (Etats-Unis, Japon, Allemagne, Royaume-Uni, Espagne, Italie). Cette préconisation devra envisager le recours à des capacités de calculs localisées hors de France, et intégrer la politique de l'Union européenne dans ce domaine (6^{ième} et 7^{ième} PCRD) ;
- Une analyse des investissements découlant de cette stratégie, conduisant à un « plan d'affaires » pluriannuel ;
- Un plan de travail sur 6 mois afin de mettre en œuvre vos recommandations.

Ces travaux me seront présentés pour approbation. Une fois validés, le groupe de travail sera chargé de mettre en place le plan de travail.

Liste des personnes rencontrées

Académie des technologies

Paul Caseau

Bull

Gérard Roucairol

CEA

Bernard Bigot
Jacques David
Alain Dechatre
Jean-Pierre Genin*⁵¹
François Gounand
Pierre Leca
François Robin
Jean-Claude Sabatier

CERFACS

Jean-Claude André

CINES

François Daumas
Thierry Porcher*

CNRS

François Etienne
Sylvie Joussaume*
Bernard Larrouturou
Antoine Petit
Serge Petiton*

⁵¹ le signe * indique les membres du groupe de travail animé par les rapporteurs.

Conférence des présidents d'université

Eric Espéret

EADS

Eric Duceau

EDF

Yves Bamberger
Jean-François Hamelin*

IBM

Luigi Brochard
Gilles Lesage
Eric Petit
Michel Teyssède

IDRIS

Victor Alessandrini*

IN2P3

Denis Linglin*

INRIA

Gilles Kahn
Stéphane Lanteri*

INSERM

Hervé Douchin
Florent Soubrier

INSU

Patrick Mascart

INTEL

Marc Dollfus

IPSL

Pascale Braconnot
Marie-Alice Foujols*
Claire Lévy
Olivier Marti

LPT Orsay

Olivier Pène

Mercator

Philippe Bahurel

Météo France

Alain Beuraud
Jean-Pierre Beysson
Eric Brun
Alain Ratier

Ministère de l'Education nationale, de l'enseignement supérieur et de la recherche

Claude Jolly
Jean-Marc Monteil

Ministère délégué à la Recherche

Michel Bidoit
Jean-Michel Dion
Jean-Jacques Gagnepain
Elisabeth Giacobino
Frédéric Gruet*
Alain Kavenoky*
Alain Lichnewsky*
Antoine Masson

Ministère délégué à l'Industrie

Emmanuel Neuville*

Laurent Rojey*

ORAP

Claudine Schmidt-Lainé

RENATER

Alain Quéré*

Dany Vandromme*

SGI

Guy Chesnot

Benoît Hallez

Hervé Oheix

Jean-Pierre Panziera

Michael Woodacre



**CONSEIL GENERAL DES TECHNOLOGIES
DE L'INFORMATION**
139 rue de Bercy
75572 PARIS CEDEX 12

**INSPECTION GENERALE DE L'ADMINISTRATION
DE L'EDUCATION NATIONALE ET DE LA RECHERCHE**
107 rue de Grenelle
75357 PARIS SP 07

Annexe au rapport
sur
La politique française
dans le domaine du calcul scientifique

PROSPECTIVES

Mars 2005

Michel HÉON

Inspecteur général de l'administration
de l'éducation nationale et de la recherche

Emmanuel SARTORIUS

Ingénieur général des Télécommunications

Table des matières

Prospectives.....	3
Introduction	3
Prospective CNRS	4
Composition du groupe CNRS.....	4
Synthèse commune CNRS et CEA/DSM.....	4
Prospective CEA	9
Contributions	9
Processeurs utilisés.....	11
Prospective INRIA	14
Prospective Calcul Intensif en Biologie	15
Composition du groupe	15
Contributions	15
Prospective nanosciences et matériaux	23
Composition du groupe	23
Tableaux synthétiques	28
Formation au calcul scientifique	42
Annexe : prospective INRIA.....	44

Prospectives

Introduction

Afin de recueillir des éléments de prospective scientifique aptes à éclairer les besoins dans le domaine du calcul intensif, la mission a procédé comme suit, après avoir rencontré les responsables des organismes concernés :

- Elle a demandé au CEA et au CNRS de fournir des perspectives à jour sur le calcul intensif. Une réunion a permis de faire un point à mi-parcours et d'échanger avec les participants au groupe de prospective CNRS réunis par Sylvie Joussaume, Directrice de l'INSU
- Elle a réuni deux groupes ad hoc pour réfléchir sur des domaines horizontaux, en minimisant l'impact des structures verticales « *disciplinaires* » ou liées à la structuration entre organismes. Ces deux groupes ont travaillé sur les sujets suivants :
 1. Le calcul intensif dans les nano-sciences et les nanotechnologies
 2. Le calcul intensif dans le domaine de la biologie.

Etant donné la durée de son intervention, la mission n'a pas tenté d'obtenir une prospective aussi affinée que celle qui serait faite par un conseil scientifique agissant dans le cadre d'une mission permanente définie par le Ministre. Son but a été d'une part de fournir les éléments permettant d'étayer les conclusions et propositions du rapport, d'autre part de se rendre compte de l'intérêt d'institutionnaliser une telle pratique, comme il est fait dans de nombreux pays.

L'objectif de cette consultation n'était pas d'obtenir une évaluation de la demande totale d'heures de calcul mais de repérer les grandes questions scientifiques dont la résolution ne peut être envisagée que si les puissances de calcul sont fortement augmentées. Il a été demandé aux groupes de perspectives de cadrer leurs réflexions dans des scénarios définis a priori par la Direction de la recherche et validés par la mission.

La mission tient à remercier tous les participants pour le travail effectué en préparant les réunions et dans les synthèses. Elle souligne dans son rapport et ses recommandations l'importance de la poursuite de ces activités et de leur intégration dans le dispositif de pilotage scientifique.

Prospective CNRS

Composition du groupe CNRS

Le groupe de travail du CNRS a été animé par Mme Sylvie Joussaume, directrice de l'INSU. Il était composé de :

- | | |
|--------------------------|--|
| • Stéphane Cordier | Sciences physiques et de la matière |
| • François Etienne | IN2P3 |
| • Patrick Le Quéré | Sciences pour l'Ingénieur |
| • Stratis Manoussis | INSU |
| • Patrick Mascart | Sciences de l'Univers |
| • Olivier Pène | Sciences physiques et de la matière |
| • Serge Petiton | Sciences et technologies de l'information et de la communication |
| • Marie-Madeleine Rohmer | Sciences chimiques |
| • Jean-Claude Thierry | Sciences de la vie |

Synthèse commune CNRS et CEA/DSM

Suite à la demande des inspecteurs M. Héon et E. Sartorius, le comité d'études prospectives sur le calcul intensif du CNRS, qui vient de se mettre en place, a conduit une première analyse des besoins en matière de calcul intensif. Il ne s'agit pas là d'une étude exhaustive des besoins mais de l'expression des besoins les plus dimensionnants. Ce travail a été mené par le CNRS avec le CEA/DSM pour les applications communes entre les deux organismes.

Le calcul intensif est devenu un outil indispensable pour nombre de domaines permettant de répondre à des enjeux de société, comme le changement climatique, le développement de l'océanographie opérationnelle ou la santé, ou à des enjeux industriels comme la combustion turbulente ou la recherche sur les médicaments. Il permet également d'aborder des questions de recherche fondamentale indispensables au développement de la connaissance comme l'astrophysique, la physique des particules et de la matière.

L'analyse par domaine fait ressortir des besoins d'augmentation en matière de calcul intensif au niveau national, exploitables immédiatement par la communauté scientifique et qui permettraient de suivre le niveau des moyens disponibles par nos partenaires proches voisins en Allemagne et au Royaume-Uni. Elle montre également, pour quelques domaines, le besoin d'avoir accès à des machines de puissance supérieure, qui pourraient être partagées au niveau européen, et permettraient de collaborer avec les USA et le Japon grâce à l'accès à des moyens de calcul comparables.

La recherche, qui s'appuie sur le calcul scientifique, exprime également le besoin d'avoir accès à une panoplie de moyens de calcul : du mésocentre, aux centres nationaux voir à quelques machines au niveau européen, afin de prendre en compte les besoins qui varient selon les domaines et selon les applications.

Le comité a fait également ressortir le rôle clé joué par les mathématiques appliquées et les recherches informatiques avec la nécessité de développer une approche interdisciplinaire combinant chercheurs des applications, mathématiciens appliqués et informaticiens. Plusieurs domaines ressentent également un déficit en soutien informatique de type ingénieur que ce soit de manière ponctuelle pour développer la parallélisation de codes ou de manière soutenue au plus près des applications selon le degré de maturité des développements des codes et l'utilisation par une large communauté.

Besoins les plus dimensionnants issus de différents domaines scientifiques :

Recherches sur le climat (CNRS & CEA/DSM)

Prévoir l'évolution future du climat et répondre aux questions posées à la société sur le réchauffement climatique suite à l'augmentation des gaz à effet de serre est entièrement dépendant de l'accès à des moyens de calcul intensif. Réduire les incertitudes, quantifier la probabilité d'événements extrêmes, quantifier les puits de carbone et leur possible évolution suite au changement du climat, étudier les impacts sur les écosystèmes et la société passe par l'augmentation de la puissance de calcul par un facteur de 10 à 1 000 par rapport aux moyens actuels afin de pouvoir intégrer la complexité du système, augmenter la résolution des modèles, affiner la représentation de processus clés comme les nuages ou les aérosols.

Ce domaine nécessite à la fois une augmentation de moyens au niveau national (5 teraflops soutenus) et l'accès à une machine européenne (50 teraflops soutenus vers 2007-2008) pour certaines expériences numériques les plus poussées en résolution et en complexité. Un projet de calculateur européen pour le climat est soutenu par la communauté scientifique qui pourrait être ouvert à une communauté plus large ayant des besoins similaires. Il peut s'intégrer dans le projet de calculateur européen proposé par l'Allemagne, la France et le Royaume-Uni, si les spécificités de la communauté climat sont bien prises en compte.

Recherches en océanographie et en sciences de la Terre

L'océanographie, en accompagnement du développement de l'océanographie opérationnelle, requiert également une grande capacité de calcul. Comprendre comment fonctionne la variabilité de l'océan, les interactions d'échelle, l'action des tourbillons dans le transport de chaleur et dans les cycles biologiques, nécessite une augmentation de calcul par des facteurs de 10 à 100 et bénéficierait d'une augmentation de calcul au niveau national voire européen.

La modélisation de la dynamique interne de la Terre, la modélisation de la réponse sismique, et des effets de site, lors d'un tremblement de Terre, la modélisation de la rupture sismique et de la génération des ondes courtes associées importantes pour le génie parasismique sont également dépendants des moyens de calcul, même si cette communauté exprime le besoin d'une hiérarchie de moyens de calcul, en particulier la nécessité d'un niveau de type mésocentre.

Recherches en astrophysique (CNRS & CEA/DSM)

La simulation numérique en astrophysique est en pleine expansion grâce au développement des codes et de la puissance de calcul. Les enjeux concernent la compréhension de l'univers qui nous entoure et l'exploitation des données des télescopes sol et des mesures depuis l'espace. Accéder à des moyens de calcul plus puissants par des facteurs 10 à 1000 permettrait de simuler la formation des galaxies, des amas de galaxies, simuler des cartes virtuelles de l'univers afin d'interpréter les grandes expériences de cosmologie comme le futur satellite Planck, de comprendre la dynamique interne du Soleil, la variabilité solaire, les mécanismes physiques de l'explosion des supernovae ou la formation des étoiles.

Recherches en ingénierie

En ingénierie, la simulation numérique en ingénierie est un outil indispensable pour aborder des enjeux industriels dans le domaine des transports, de l'énergie et du développement durable. Réduire la consommation des moteurs, réduire les émissions de gaz à effet de serre, réduire le bruit sont autant d'enjeux sociétaux et industriels qui nécessitent de disposer de capacités prédictives améliorées qui reposent sur la modélisation et simulation numérique intensive. Simulation de la turbulence, des réactions chimiques dans des fluides, des transferts dans des milieux hétérogènes, les interactions entre fluides et matériaux, la réponse des matériaux à différentes contraintes, le développement de nouveaux matériaux s'appuient sur la simulation numérique au sein de projets alliant équipes de recherche en interaction étroite avec les industriels concernés.

Ces domaines d'application sont actuellement limités par l'accès aux moyens de calcul disponibles. Prendre en compte la complexité des phénomènes, celle additionnelle due à leurs couplages, la gamme des échelles spatiales et temporelles impliquées, le grand nombre de réalisations pour tester différents cas et permettre leur optimisation ou leur contrôle, nécessitent une augmentation des moyens de calcul disponibles que ce soit au niveau national et au niveau européen. Pour prendre un des grands défis scientifiques de l'ingénierie scientifique, le problème de la combustion turbulente dans des moteurs, une augmentation par un facteur 10 à 100 permettrait une meilleure description de la turbulence et ainsi d'identifier des zones de gaz frais et de gaz chauds dont le comportement peut être notablement différent que ce soit en terme de turbulence, d'émissions radiatives ou de formation de polluants.

Recherches en physique fondamentale (CNRS & CEA/DSM)

Comprendre les lois fondamentales de la physique, la composition de la matière, des particules à la composition du cosmos, requiert l'accès à des moyens de calcul. Deux domaines sont

particulièrement demandeurs : la physique des particules en accompagnement des expériences de physique des particules et la chromodynamique quantique.

En physique des particules, les besoins dimensionnants en matière de calcul concernent l'accompagnement des expériences du *Large Hadron Collider* avec des enjeux comme la validation du modèle standard, la prospection de particules supersymétriques, la compréhension de l'absence d'antimatière dans le cosmos ou même l'immense quête de la supposée matière noire galactique. Cette communauté a des besoins très spécifiques liés à la très grande masse de données à traiter qui se compte en petaoctets. Afin d'exploiter au mieux toutes ces données, ils développent une infrastructure répartie au niveau mondial et contribuent fortement au développement de grilles de calcul et grilles de données qui répondent parfaitement à leurs besoins. Ces développements, au départ pour la physique fondamentale, devraient à terme pouvoir servir à de nombreuses autres applications en particulier pour la société et l'industrie.

La chromodynamique quantique s'attache à comprendre l'interaction forte qui lie ensemble quarks et gluons au sein des protons et neutrons. La France a participé à un projet de R & D de l'ordinateur européen apeNEXT qui est adapté aux besoins de cette discipline. L'enjeu pour cette communauté de plus petite taille que la physique des particules serait de pouvoir contribuer à l'implémentation d'une telle machine au niveau européen de 10 teraflops soutenus dans l'immédiat, 1 petaflop vers 2010. Ces développements devraient avoir des retombées dans l'industrie des ordinateurs.

Recherches en chimie

La chimie est un des domaines utilisateur du calcul intensif que ce soit pour des applications industrielles ou des enjeux de société comme l'industrie pharmaceutique. A titre d'exemple dimensionnant on peut citer l'utilisation des méthodes de la chimie théorique pour étudier les structures, les propriétés et la réactivité des complexes de métaux de transition. Il s'agit de comprendre la spécificité des réactions chimiques et le rôle joué dans une structure et dans une réaction par tous les groupes chimiques constituant les molécules. Ces méthodes ont des applications dans la catalyse ou dans la diffusion des ions lithium dans des batteries. Ils nécessitent de représenter des systèmes de grande taille avec des calculs numériques itératifs très demandeurs en temps de calcul.

Dans le domaine des médicaments, un autre exemple est cité ci-dessous avec les applications biologiques : celui des nanostructures biologiques à l'interface entre chimie et biologie.

Recherches en biologie

La biologie est un domaine en émergence dans le domaine du calcul scientifique. Les besoins concernent tout aussi bien la fabrication de médicaments, avec en particulier l'ingénierie des ligands, la compréhension de mécanismes réactionnels, le repliement de protéines, la simulation de la complexité du vivant, comme les cellules ou les écosystèmes. Un besoin particulier d'accès interactif à de grandes bases de données en réseau est un également un enjeu majeur pour la biologie.

A titre d'exemple dimensionnant on peut citer le domaine à l'interface entre la biologie et la chimie : la simulation de nanostructures biologiques. Celle-ci permet d'étudier l'assemblage et la transformation de macromolécules biologiques, des réactions enzymologiques, à la clé de la fabrication de médicaments ou d'étude de mécanismes biologiques. Le repliement inverse des protéines est également un domaine qui requiert une augmentation de puissance de calcul. L'évolution rapide des besoins de calcul intensif dans ce domaine mérite une analyse approfondie.

Recherches en mathématiques appliquées et en informatique

Recherches en mathématiques appliquées et en informatique jouent un rôle particulier vis-à-vis de l'utilisation du calcul intensif.

Mathématiciens experts en analyse numérique, en modélisation et en algorithmique jouent un rôle important dans un très grand nombre de projets de calcul scientifique. On peut citer à titre d'exemple l'apport des mathématiques dans le développement de modèles du stockage géologique profond des déchets nucléaires. L'algorithmie joue également un rôle clé qui s'amplifiera dans les années à venir à mesure que les progrès du calcul intensif devront reposer de plus en plus sur l'inventivité algorithmique à mesure que les progrès des processeurs dus simplement à des gravures fines vont se tarir.

De nombreuses recherches en STIC peuvent également avoir un impact sur le calcul scientifique intensif. Certaines de ces recherches se font principalement en amont : sur les architectures, la compilation, la programmation, d'autres sont induites par les applications dimensionnantes comme la visualisation, l'analyse de données. Enfin certaines peuvent demander une collaboration étroite avec les applications : stabilité numérique, arithmétique, algorithmique, complexité, calcul formel. Si la communauté STIC calcule peu sur les centres nationaux, les recherches concernées ont souvent besoin pour effectuer des mesures précises d'avoir accès aux machines en mode dédié. Il est aussi nécessaire de pouvoir avoir accès à différents types de machines et des plates-formes comme Grille 5000.

Prospective CEA

La prospective CEA est issue pour l'essentiel de l'analyse des besoins en calcul scientifique intensif lancée par le Comité directeur de l'informatique scientifique du CEA à la mi 2004. Répondant à la demande de la mission de disposer d'éléments de prospective selon les disciplines, la prospective pour l'astrophysique, la climatologie et la physique à haute énergie a été élaborée avec le CNRS et incluse dans la synthèse commune CNRS et CEA/DSM. S'agissant des domaines « nano-sciences et nanotechnologies » et « biologie », des spécialistes du CEA ont participé aux deux groupes ad hoc réunis par la mission. Enfin les éléments de prospective au CEA sur le calcul intensif dans le domaine de l'énergie nucléaire (fission et fusion) sont fournis ci-dessous.

Contributions

Besoins de calcul intensif de la Direction de l'Energie Nucléaire contribution de la Direction de l'Energie Nucléaire du CEA.
--

Introduction

Toutes les disciplines de la physique qui entrent en jeu dans le cycle de vie des installations nucléaires de la conception au démantèlement et au stockage des déchets en passant par le fonctionnement normal ou accidentel utilisent des outils de simulation. Il s'agit principalement des matériaux, de la neutronique, de la thermohydraulique, du combustible, du stockage géologique et de la mécanique.

Ces domaines de simulation ne sont évidemment pas indépendants, et la plupart des phénomènes à simuler nécessitent des couplages entre les différentes disciplines et de plus en plus souvent des couplages d'échelles dans une même discipline.

Trois disciplines dominent aujourd'hui et continueront dominer demain les moyens de calcul centralisé : les **matériaux**, la **neutronique** et la **thermohydraulique**. Les autres disciplines qui nécessitent beaucoup moins de puissance de calcul, relèvent davantage de ressources locales que de grands moyens mutualisés, elles ont donc volontairement été exclues de la présente analyse prospective de besoin.

1.1/ Comportement des matériaux sous irradiation

L'accumulation de défauts d'irradiation peut entraîner une modification de la microstructure des matériaux et avoir des effets sur les propriétés macroscopiques, la tenue mécanique par exemple. Cette problématique concerne :

- la prolongation de la durée de vie des installations nucléaires,
- le stockage géologique,
- L'optimisation des combustibles actuels et la conception de combustibles pour les réacteurs futurs...

Les simulations actuelles représentent une centaine d'atomes, ce qui est suffisant pour la représentation de **défauts** ou **d'amas de défauts isolés**. Elles ont permis le calcul d'un grand nombre de configurations d'amas de défauts dans le fer. Il faut maintenant :

- modéliser des matériaux plus complexes : Fe(C), Fe(Cu).
- modéliser de défauts plus étendus (doublement du nombre actuel d'atomes - multiplication par 4 des ressources de calcul)
- modéliser des dislocations (quadruplement du nombre actuel d'atomes - multiplication par 16 des ressources de calcul)
- automatiser les couplages 2 échelles pour permettre par exemple : des couplages *ab initio* – dynamique moléculaire (pour l'ajustement des potentiels) ; des couplages dynamique moléculaire - Monte Carlo (pour la description de défaut d'irradiation produit par des cascades).
- à plus long terme généraliser le couplage des trois échelles (*ab initio*, dynamique moléculaire et Monte-Carlo).

1.2/ Physique des réacteurs

Cette problématique concerne en particulier :

- la conception et de développement de cœurs nucléaires nouveaux,
- l'optimisation du fonctionnement des réacteurs actuels,
- la sûreté des installations et le démantèlement...

Parmi les principales évolutions prévisibles qui consommeront de grosses ressources de calcul on peut citer :

- Les couplages neutronique – photonique pour la conception et le développement de futurs réacteurs expérimentaux.
- Les couplages neutronique probabiliste – combustible qui devraient rester limitée jusqu'en 2007 puis croître rapidement dans les années suivantes.
- Les calculs 3D fins de cœur d'un cycle complet de type cellule par cellule
- Les calculs de radioprotection d'une installation complète sans approximations et les calculs d'activation des structures pour des applications de démantèlement.

1.3/ Thermohydraulique

Dans un horizon de 3 à 4 ans les besoins en thermohydraulique vont changer brutalement d'ordre de grandeur par la généralisation de calculs de Simulation Numérique Directe. Ces simulations sont et seront de plus en plus utilisées en support à la modélisation physique des échelles macroscopiques en particulier lorsque le support expérimental n'est plus accessible ou trop onéreux.

2/ Quantification des besoins

La présente analyse est limitée aux besoins en ressources de calcul scalaire. Les besoins vectoriels ne sont pas abordés, l'analyse faite montrant une relative stabilité dans l'utilisation de ce type de ressource.

teraflops crête scalaire	2004	2007-2009	2009-2012
Matériaux	0,50	2,00	4,00
Thermohydraulique	0,25	1,00	4,00
Neutronique	0,25	1,00	2,00
Total	1,00	4,00	10,00

Processeurs utilisés	2004	2007	2009
Matériaux	30	60	200
Thermohydraulique	30	60	500
Neutronique	20	60	100

2/ Positionnement par rapport aux scénarios envisagés

L'analyse présentée ci-dessus montre un quadruplement des besoins de puissance de calcul de la DEN entre fin 2004 et début 2007. Le scénario A ne peut d'évidence pas répondre au besoin de disposer de 4 teraflop/s dès 2007 (la DEN consommerait à elle seule 16% des ressources ce qui ne paraît pas envisageable). A court-moyen terme seul le scénario B permettrait de répondre aux besoins de la DEN. A moyen-long terme les 3 scénarios B, C et D répondent à nos besoins en terme de puissance brute (10 teraflop/s dès 2009). La discrimination devra se faire sur d'autres critères en particulier sur le mode d'utilisation. Deux modes sont envisagés :

- Mode « Usage par une communauté » : Utilisation d'au plus 50% de la configuration par un utilisateur à un instant donné.
- Mode « Grand Challenge » : Utilisation d'au moins 50% de la configuration pendant une durée importante. Seul mode disponible dans les scénarios C et D.

Si les besoins en neutronique relèvent du premier mode d'utilisation, les besoins en matériaux et en thermohydraulique relèvent davantage du mode « grand challenge ».

Conclusion

Dans ce qui précède nous avons limité l'analyse au seul niveau de puissance. D'autres aspects sont cependant très importants pour le dimensionnement de moyens de calcul :

- la taille de la mémoire proche (disponible sur une même carte et partagée par un nombre limité de processeurs). L'expérience montre qu'un équilibre doit être recherché entre la taille de cette mémoire et la puissance du processeur. *Augmenter d'un facteur x la puissance des processeurs nécessite, pour conserver cet équilibre, d'augmenter la taille de la mémoire proche du même facteur ;*
- les temps d'accès de la mémoire éloignée (disponible sur d'autres cartes et partagée par un grand nombre de processeurs). Compte tenu des méthodes de parallélisation utilisées il est primordial que ces temps d'accès soient les plus petits possibles. *Les caractéristiques des réseaux d'interconnexion doivent donc être un élément fort de choix.*
- la disponibilité de moyens de pré et post traitement...

Des besoins de visualisation 3D, de type stéréoscopique, existent par exemple pour suivre la migration d'atomes ; la visualisation des chemins de migration ; etc. La mise à disposition de ce type d'outils serait un plus incontestable, pour voir et éventuellement agir durant un calcul.

Au-delà d'une simple augmentation de puissance, de mémoire et de débits d'entrées sorties, certains des nouveaux besoins, tout particulièrement les couplages de disciplines et d'échelles, vont imposer des choix différents dans : les méthodes de gestion des ressources ; les outils disponibles sur les machines centrales ; les règles de sécurité ; le calcul conjoint sur moyens locaux et centraux, etc.

<p>Le calcul dans le cadre du projet ITER contribution de J. P. Génin, CEA</p>

Le défi scientifique et technique posé à la communauté internationale par la construction du réacteur expérimental de fusion thermonucléaire par confinement magnétique ITER demande un travail intense au niveau de la théorie et de la modélisation, pour faire avancer non seulement la compréhension des phénomènes de base dans des domaines de la physique extrêmement variés, mais

également pour intégrer l'ensemble de ces modèles dans des codes de calcul capables de prédire avec une fiabilité suffisante le comportement du plasma et des divers composants du réacteur : les « simulateurs tokamak ». Il s'agit d'assurer à la fois le succès de l'expérience et sa sûreté.

ITER représente un véritable tournant pour la modélisation, qui doit désormais réaliser cette intégration à l'horizon 2015-2020. Or les recherches actuelles montrent que les processus physiques qui régissent, par exemple, le comportement du cœur du plasma impliquent de résoudre les équations cinétiques complètes dépendantes du temps (2 ou 3 dimensions dans l'espace des vitesses dans une géométrie principalement à deux dimensions), couplées aux équations de l'électromagnétisme. Les échelles de temps en jeu couvrent 4 à 5 ordres de grandeurs (magnétohydrodynamique, transport de la chaleur et des particules, évolution du profil de courant plasma). L'intégration de cette physique du cœur du plasma aux phénomènes périphériques dits d'interaction plasma-paroi, eux-mêmes influencés par la physico-chimie locale dans une géométrie à trois dimensions spatiales avec de nouvelles échelles de temps requiert des moyens de calcul centralisé hors de la portée actuelle de la communauté fusion magnétique.

La nécessité du travail est reconnue au niveau mondial depuis deux à trois ans, principalement par les partenaires Etats-Unis (Fusion Simulator Project (FSP), ou initiative SciDAC), et Europe (European Task Force on Integrated Tokamak Modelling). Il faut noter que le Japon a également débuté une réflexion dans ce sens depuis quelques mois. Ces initiatives de modélisation intégrée vont cependant se heurter rapidement au problème des ressources de calcul centralisé. Toute augmentation de la puissance de calcul disponible permet à la fois d'affiner les modèles, mais également de les coupler entre eux de manière à gagner sur la cohérence globale. Le projet américain FSP a quantifié ce que l'utilisation des « supercalculateurs » actuels peut apporter. On peut citer quelques chiffres : au niveau des ressources de calcul pur certaines simulations de magnétohydrodynamique des plasmas en ignition approchent la centaine de teraflops. La résolution de certains problèmes de propagation/absorption d'ondes de chauffage additionnel requiert un nombre d'opérations d'environ un millier de teraflops*heure. La résolution de la dynamique des électrons sur l'échelle du temps de confinement en régime turbulent nécessite au moins 10 000 teraflops*heure par simulation. Il faut également souligner les besoins concomitants en matière de stockage de données et de rapidité des réseaux d'échanges liés aux quantités de données traitées. Le projet FSP souhaite disposer de 4 teraflops maintenant, de 10 dans un à trois ans et de 100 au-delà de trois ans.

En résumé, une simulation typique de MHD ou de turbulence nécessitera 103 teraflops*heure en 2008 et 10 fois plus (104 teraflops*heure) à l'horizon 2012 (pour un démarrage prévu d'ITER en 2015), lorsque la physique complète des électrons sera incluse. L'objectif est d'optimiser l'exploitation d'ITER, i.e. de préparer les expériences au mieux par des simulations les plus détaillées possible de la stabilité, du confinement, du chauffage et de l'extraction de puissance des plasmas de fusion. Un projet scientifique, par exemple la préparation d'un scénario dans ITER, requiert un nombre élevé de simulations (de l'ordre de la centaine). L'estimation est donc de l'ordre de 10^5 teraflops*heure à l'horizon 2008 et 10^6 teraflops*heure à l'horizon 2012 pour chaque projet scientifique.

Prospective INRIA

Voir Annexe.

Prospective Calcul Intensif en Biologie

Composition du groupe

Vincent Breton	CNRS / IN2P3
Christophe Combet	CNRS / IBCP
Gilbert Deleage	CNRS / IBCP
Martin Field	CEA
Christine Gaspin	INRA / LBIA
Nicolas Jacq	CNRS / IN2P3 LPC Clermont Ferrand
Richard Lavery	IBPC Paris
Michael Nilges	Institut Pasteur
François Rodolphe	INRA Jouy-en-Josas
William Saurin	Genomining
Thomas Simonson	Ecole Polytechnique
Jean-Claude Thierry	CNRS SdV / IGBMC Strasbourg

Contributions

**Quelques enjeux majeurs en biologie
computationnelle
contribution de MM. R. Lavery et T. Simonson**

Le problème du repliement inverse

Avec l'explosion de l'information génomique, l'écart entre le nombre de séquences et le nombre de structures de protéines connues continue de se creuser. La simulation moléculaire peut aider à résoudre ce problème grâce aux techniques dites de *repliement inverse*. Il s'agit d'identifier les séquences les plus favorables correspondant à une structure tridimensionnelle donnée. Un enjeu actuel est d'appliquer la méthode aux 3000 repliements protéiques connus, en utilisant par exemple l'évolution dirigée *in silico*, qui a atteint un niveau raisonnable de maturité et continue de progresser rapidement. On peut estimer à 5-10 ans le temps monoprocesseur nécessaire avec les modèles actuels. Pour la prochaine génération de modèles théoriques, aujourd'hui en développement, il faudra augmenter la puissance de calcul par 10. La "cartographie des séquences" obtenue permettra d'avancer dans l'attribution des structures (e. g. dans le contexte de la génomique structurale), l'attribution des fonctions (pour l'annotation des génomes et la protéomique), et l'ingénierie de protéines avec des fonctions nouvelles.

La reconnaissance protéine-ligand et l'ingénierie de ligands : l'apport de simulations quantitatives

L'ingénierie de ligands s'appuie aujourd'hui sur un spectre de méthodes de coûts et fiabilités variables. Dans les sociétés pharmaceutiques, des méthodes simples sont employées, mais également des méthodes assez sophistiquées, et une demande existe pour des méthodes véritablement quantitatives. En recherche fondamentale, des méthodes coûteuses mais potentiellement très fiables sont développées depuis dix ans. Les progrès en puissance de calcul et une expérience grandissante sont en train de porter à maturité ces méthodes "exactes", dites "calculs d'énergie libre". Pour augmenter leur fiabilité, il faut pouvoir multiplier les simulations de dynamique moléculaire et aussi explorer de nouveaux degrés de liberté, comme le couplage entre fixation de ligand et fixation/libération de protons. Pour appliquer ces méthodes efficacement, nous devons gagner un ordre de grandeur en puissance de calcul. Pour un seul projet (= un chercheur), il faut une puissance dédiée qui équivaut à une grappe de 32 processeurs Intel de dernière génération (= 300 000 heures CPU au CINES). De plus, il faut pouvoir accéder à des pointes de vitesse 2 ou 3 fois supérieures lors des phases critiques d'un projet.

Assemblage de complexes protéiques

La plupart des processus biologiques impliquent l'assemblage de complexes protéiques. Notre connaissance de la gamme de ses assemblages est en train d'exploser grâce à de nouvelles techniques expérimentales. Ces techniques témoignent de l'existence de milliers d'interactions binaires au sein des cellules, et permettent de dresser les premières cartes de "l'interactome". Ce domaine nécessite néanmoins de nouvelles collaborations entre bioinformaticiens et structuralistes pour traduire des informations binaires en modèles moléculaires, permettant ainsi l'obtention des données structurales, thermodynamiques et cinétiques pour des complexes bi- ou multimoléculaires. Grâce à de tels modèles il devient également possible d'étudier la perturbation des complexes suite aux mutations ou aux interactions avec des agents externes. Ce travail implique des efforts significatifs pour améliorer les techniques existantes pour prédire la conformation des complexes protéine-protéine. Il faut notamment pouvoir tenir compte de la flexibilité intrinsèque des protéines et savoir intégrer des informations fournies par l'étude de l'évolution des séquences. La quantité d'information à traiter dans ce domaine doit stimuler l'exploitation des ressources en calcul distribué.

Contrôle de l'expression génétique

Une meilleure compréhension du contrôle de l'expression génétique reste un enjeu majeur pour la biologie. Les zones de fixation des protéines impliquées dans ce contrôle sont plus difficiles à déceler que les gènes correspondants, notamment à cause des séquences consensus mal définies et d'un manque d'information sur l'ensemble des protéines impliquées. Dans le cas des eucaryotes, il faut ajouter la complexité liée à l'empaquetage du génome et notamment au positionnement des nucléosomes. Les études de modélisation peuvent contribuer à résoudre ces problèmes en permettant d'accéder aux facteurs physiques qui sous-tendent les interactions protéine-ADN. De telles données peuvent être exploitées en parallèle avec des informations provenant de l'étude des séquences pour améliorer les capacités de détection, mais aussi peuvent servir pour aborder la modélisation de protéines homologues pour lesquelles il n'y a pas d'information sur leurs interactions. Comme dans la

plupart des domaines liés à l'analyse des génomes, une étude de la multiplicité des éléments en interaction passe par l'exploitation des moyens de calcul significatifs, mais aussi par l'intégration des informations dans des modèles à différents niveaux qui nécessiteront des collaborations transdisciplinaires efficaces.

<p style="text-align: center;">Exploitation des données produites par la biologie contribution de C. Gaspin</p>
--

Les volumes considérables de données produites par la biologie interpellent les mathématiques et l'informatique autour de leur exploitation. Deux familles de problèmes pourraient bénéficier d'une structure offrant des possibilités en calcul intensif. La première famille de problèmes s'appuie sur des algorithmes dont la complexité est « acceptable », typiquement, des algorithmes linéaires ou de faible complexité polynomiale. En général, le volume considérable des données (le volume des données croit de manière exponentielle) à analyser constitue le frein majeur (voir par exemple 1)). La deuxième famille contient les problèmes pour lesquels les algorithmes sont coûteux voire NP-durs. (voir par exemple 2).

Comparaisons de séquences (1)

Les séquences complètes des premiers génomes séquencés ont donné la possibilité d'accéder à des organisations structurales « statiques » des génomes. Le volume déjà considérable des données a très rapidement imposé, dès les premiers lots de séquences disponibles, l'utilisation à grande échelle d'algorithmes efficaces (souvent des heuristiques) pour la comparaison des séquences entre elles (le logiciel BLAST en est un exemple). C'est donc ici la distribution d'une masse considérable de calculs de comparaisons sur un grand nombre de processeurs qui pourra constituer une avancée et donner un avantage à une équipe à un instant donné (ex : assemblage du génome humain). Les enjeux sont par exemple : une prise de position dans un contexte/programme international d'assemblage/annotation avec à la clef, un accès privilégié aux premières séquences annotées. Des architectures spécialisées existent aussi pour ce type de calculs.

Alignement de séquences et phylogénie (2)

Les premiers algorithmes exacts d'alignement de séquences (programmation dynamique), mais aussi les méthodes d'analyse de données, de maximum de vraisemblances, les modèles markoviens et autres... sont des algorithmes dont la complexité (parfois exponentielle) devient un frein prohibitif dès que la taille du problème croit (ex : longueur d'une séquence, nombre de séquences, d'individus...). La reconstruction phylogénétique constitue une illustration de ce type de problème. Une phylogénie représente l'histoire évolutive d'un groupe de gène ou de taxa. Les algorithmes largement utilisés de nos jours s'appuient sur des modèles basés sur l'évaluation des substitutions (mutations) dans les données de séquences. Ils peuvent se révéler très coûteux selon les méthodes utilisées (méthode de parcimonie et de maximum de vraisemblance) et pourraient bénéficier

pleinement d'un centre ouvert de calcul intensif (cf. ¹). On peut citer parmi les enjeux, hors une meilleure connaissance de l'histoire évolutive, une meilleure compréhension des maladies et de leur mode de propagation. Avec l'accès aux génomes complets et à leur annotation, la reconstruction phylogénétique invente aujourd'hui de nouveaux modèles qui tentent de reconstruire l'histoire des génomes à partir de leur organisation structurale et d'un ensemble d'opérations autorisées (perte, duplication, réarrangement, ...) sur les gènes. Les algorithmes restent cependant très combinatoires et pourraient aussi pleinement bénéficier de telles infrastructures. D'autres questions sont aussi concernées : analyse des données d'expressions, recherche de motifs, alignement multiple, cartographie des marqueurs moléculaires... Certaines de ces questions sont déjà ouvertes devraient s'ouvrir à des modèles et méthodes considérant les aspects quantitatifs (analyse des données d'expression notamment – transcriptome mais aussi protéome).

**Portails intégrés d'analyses informatiques pour la biologie
contribution de C. Combet, laboratoire PBIL-IBCP**

A l'ère de la biologie à grande échelle, la quantité toujours croissante de données biologiques hétérogènes nécessite de plus en plus de traitements informatiques pour le stockage et l'exploitation de celles-ci. Devant le grand nombre de données et d'outils de traitements de celles-ci, il est nécessaire de mettre en place pour les utilisateurs finaux que sont les biologistes des services intégrés, conviviaux et interactifs d'analyses, de modélisation et de simulation des macromolécules biologiques qui les aident dans leurs recherches au quotidien. De tels services devront être accessibles par Internet et proposer une interface simple et convivial aux utilisateurs en intégrant dans un enchaînement logique les outils de traitements des données avec un temps de restitution le plus court possible (ce qui nécessite des moyens de calculs importants). Les problématiques adressées dans la mise en place de tels portails sont l'uniformisation des données, la définition de protocoles d'analyses, la "parallélisation" des calculs, l'intégration et la définition des "interfaces homme-machines". La faisabilité et l'utilité de tels portails ont été démontrées au travers de nombreux portails mis en place par les laboratoires nationaux de bioinformatique sur leurs ressources propres et qui sont plus ou moins limités dans leur capacité pour des raisons de temps de calculs.

**Calcul intensif et mésocentre en biologie
contribution de V. Breton et N. Jacq, CNRS/LPC**

Note liminaire : *La mission a décidé d'intégrer ici cette contribution qui traite de l'organisation requise de la recherche en biologie et bioinformatique pour prendre en compte l'intervention de plus en plus fondamentale de l'information sous des formes extrêmement massives. Ceci faisait partie des questions qui ont été posées au groupe de travail ad hoc. Toutefois, dans ses*

¹ Article de D.A. Bader, "Computational Biology and High Performance Computing", Comm. ACM, Vol 47, N° 11, Nov 2004

conclusions, la mission a estimé préférable de ne pas aborder cette question, qui a son avis doit faire l'objet d'une étude plus complète de la part le ministère en charge de la Recherche.

Introduction

L'objectif de cette note est de présenter une réflexion sur la mise à disposition des biologistes de ressources de calcul et de stockage des données ainsi que de services bioinformatiques intégrant des compétences indissociables en biologie et en informatique, pour la création et la maintenance d'outils informatiques et de banques de données, le traitement de données pour la génération de connaissance, et la formation.

Afin d'en garantir la pertinence biologique, ces infrastructures seront adossées à un ou plusieurs laboratoires de biologie de visibilité nationale et internationale.

Perspective sur les besoins en calcul scientifique et en traitement de données

Les besoins exprimés par les biologistes peuvent se résumer en quelques points essentiels :

- Disposer ou développer des bases de données généralistes ou thématiques, validées, nécessaires à leur recherche, mises à jour en temps réel dans un environnement convivial,
- Disposer des algorithmes et procédures adaptés à une recherche en constante évolution dans un environnement convivial et les faire évoluer en permanence,
- Disposer de ressources de calcul mais surtout humaines suffisantes pour un travail d'annotation et d'analyse,
- Pouvoir combiner complexité informatique et innovation biologique sous-tendue par les demandes exprimées précédemment, les outils génériques nécessitant de plus en plus une utilisation intégrant la complexité spécifique du domaine.

Pour répondre à ce cahier des charges, les centres européens et nationaux de ressource en bioinformatique ont notamment développé des systèmes de mise à jour et d'accès aux bases de données, des boîtes à outils d'algorithmes et des portails web mettant ces outils à la disposition des biologistes via Internet dans un environnement simple et raisonnablement convivial.

Ces portails se sont avérés des moyens puissants pour permettre l'analyse des données biologiques mais ils ont été confrontés à plusieurs défis :

- la complexité et la diversité (séquences, images, données de puces, CGH arrays, données structurales, données bibliographiques, criblages pour le médicament, ...) de l'information biologique qui ont suscité une multiplication des bases de données et des algorithmes,
- Le volume exponentiellement croissant des données produites qui a induit une augmentation tout aussi rapide des besoins de calcul, de stockage et de mise à jour des données,
- La multiplication des biologistes produisant et analysant des données de la génomique et de la post-génomique qui a créé un phénomène de goulot d'étranglement dans l'accès aux portails et aux ressources offertes par ces portails,

- Surtout la non-adaptation de ces sites à la spécificité des requêtes et à l'émergence de problèmes multifactoriels impliquant la combinaison de données de natures complémentaires (séquences et puces par exemple) nécessitant autant de nouveaux outils (clustering), conditions incontournables à l'exploitation des données.

C'est ainsi qu'à côté du centre national français Infobiogen, des portails thématiques (génomique, protéomique, ...) ont vu le jour dans plusieurs régions françaises notamment dans le cadre des génopôles. Ces portails régionaux ont permis d'absorber une fraction significative des besoins des biologistes pour certains dans un contexte local, régional et international. Cependant, ces portails régionaux ont leur propre architecture, assurent eux-mêmes la mise à jour des bases de données et la croissance de ces bases induit une pression croissante en ressource humaine pour la maintenance du portail.

Nous observons depuis 4 ans dans la communauté des biologistes une approche très pragmatique : ils vont « à la pêche » sur Internet des sites ou des informations qu'ils peuvent exploiter, s'adaptant presque au jour le jour aux temps de réponses offerts par les différents portails qu'ils connaissent.

La situation actuelle laisse ainsi apparaître d'importantes lacunes :

- la politique nationale est seulement en cours de mise en place pour la coordination de la gestion et de la maintenance de ces portails donnant accès à des ressources bioinformatiques dans l'hexagone. Cette action est actuellement coordonnée et des réunions de travail planifiées en 2005 par F. Soubriet et C. Gautier en coordination avec les organismes dont le CNRS (J.C. Thierry).
- Chaque portail fait sa propre recherche de fond et fait évoluer son offre en calcul ou en données en fonction des financements obtenus, de ses propres thèmes de recherche et de ses ressources humaines.

Dans ce contexte, accumuler des teraflops de puissance de calcul n'aura qu'une utilité limitée si ces teraflops ne sont pas proprement interfacés à des services bioinformatiques accessibles de façon conviviale. La croissance exponentielle des données, la multiplication des utilisateurs de ces données et la diversité des compétences requises rend le modèle centralisé de plus en plus difficile à gérer. Les nouvelles technologies, notamment les grilles, permettent de déporter les calculs les plus gourmands en teraflops vers des centres de calcul spécialisés tels que l'IDRIS, le CINES ou ceux de l'IN2P3. Gérer un centre de calcul est un métier différent de celui de gérer un centre de ressources en bioinformatique, qui doit maintenant comme l'EBI ou le NCBI combiner expériences biologiques et informatiques (algorithmique, mathématiques appliquées,...).

Les clefs d'une politique de dynamisation de la recherche en biologie en France nous apparaissent donc les suivantes à travers une décentralisation de l'offre de calculs :

- Mise en réseau de portails offrant de **façons concertées et complémentaires** des services communs et des services spécifiques,
- **mise en place d'une politique commune de gestion de la mise à jour des bases de données** à l'ensemble des prestataires de services,
- généralisation des **accords entre centres de ressources en bioinformatique et centres de calcul intensif et à haute performance** pour que ces derniers soient de plus en plus prestataires de ressources de calcul et de stockage en toute transparence pour l'utilisateur,
- Evolution des centres de calculs intensifs vers une réponse temps réels pour l'indexation des banques (quotidienne) et les requêtes des biologistes.

Propositions d'action

Nous proposons donc de mettre en place un modèle hiérarchique du calcul pour la biologie.

- Infobiogen ou tout autre centre national intéressé (*niveau 0*) pourrait par exemple être le centre de gravité de ce modèle hiérarchique. Une politique commune de gestion des mises à jour des bases de données pourrait ainsi être mise en place et coordonnée par le centre de niveau 0 pour en réduire le coût humain.
- Autour du portail national, seraient créés des "grands si possible" *centres régionaux (niveau 1)* disposant de ressources humaines pour la mise à jour des données et mettant à disposition un ou des portails web pour les utilisateurs. Ces centres de niveau 1 disposeraient de ressources de calcul et de stockage significatives et seraient reliés par des réseaux à très haut débit à des centres de calculs (centres nationaux, mésocentres) fournissant des ressources beaucoup plus importantes de calcul et de stockage.
- Autour de ces grands centres régionaux seraient mis en place des centres régionaux plus petits (*niveau 2*) disposant de ressources plus réduites pour mettre à disposition des services plus limités mais reliés aux centres de niveau 1 et de niveau 0 pour les services au-delà de leurs propres capacités.

Un tel modèle intègre la démarche actuelle du Ministère dans le cadre des génopôles. Les candidats naturels pour les centres de niveau 1 seront les centres de ressources bioinformatiques des génopôles. Cependant, ces centres doivent identifier des centres de calcul avec lesquels ils puissent collaborer selon le modèle proposé. Une telle approche hiérarchique permet de mieux absorber la communauté d'utilisateurs et privilégie les ressources de proximité.

Des synergies fortes pourraient être développées avec l'EBI. Un bon exemple de synergie avec l'EBI est le réseau d'excellence Embrace sur les grilles pour la bioinformatique (<http://www2.cnrs.fr/presse/communique/610.htm>). Son but est de standardiser l'accès aux innombrables données issues des projets de génomique, afin que les chercheurs puissent les consulter et les exploiter facilement. Le CNRS (à travers l'Institut de biologie et de chimie des protéines (IBCP) de Lyon, le centre de calcul de Lyon et le Laboratoire de physique corpusculaire (LPC) de Clermont-Ferrand) est responsable de la veille technologique : il fera en sorte qu'Embrace bénéficie des dernières avancées en matière de grille de calcul et que les choix retenus soient les mieux adaptés aux problématiques de la bioinformatique. La participation du CNRS au projet Embrace permettrait de faire bénéficier l'ensemble des centres des derniers développements (API, standards...) liés à ce réseau d'excellence, ainsi que d'un accès facilité et peu coûteux en ressources humaines à toutes les bases de données et algorithmes de ce réseau.

Dans ce contexte, le CNRS est associé au Généthon pour des premiers essais de gridification d'applications développées au niveau national, montrant ainsi son intérêt pour ce type de développement.

La technologie de grille semble être donc l'un des outils possibles pour mutualiser et partager les ressources (calcul, stockage, algorithmes, bases de données) entre les différents acteurs de ce modèle. Les fermes des centres de calcul de l'IN2P3, de l'IBCP, et bientôt du CINES, sont déjà nœuds du projet européen EGEE (<http://public.eu-egee.org>), dont le deuxième champ pilote d'applications est le biomédical.

Comme l'a demandé le Ministère à travers les actions génopôles, les centres acteurs de ce modèle doivent s'engager sur un cahier des charges précis des services offerts et doivent faire l'objet

d'une évaluation par rapport à ce cahier des charges. Ce modèle pourrait démarrer avec quelques centres pilotes.

Prospective nanosciences et matériaux

Composition du groupe

Xavier Blase	CNRS/LPMCN Lyon
Christophe Delerue	CNRS/IEMN
Thierry Deutsch	CEA/DSM/DRFMC
Patrice Hesto	IEF CNRS et Université Paris XI
François Jollet	CEA /DAM
Jean-Luc Leray	CEA/DAM
Georges Martin	CEA/HC
Lucia Reining	Ecole Polytechnique.
François Willaime	CEA/DEN
Ludger Wirtz	CNRS/IEMN
Gilles Zerah	CEA/DAM

Note : *La mission remercie Georges Martin d'avoir attiré son attention sur de nombreux aspects liés à la simulation numérique des matériaux intéressant les industries nationales, ainsi que sur le rapport du DoE sur la « Science numérique des matériaux appliquée à la fusion et aux réacteurs de 4^{ème} génération² »*

² *Workshop on Advanced Computational Materials Science : Application to Fusion and Generation IV Fission Reactors*, DoE, Washington DC, 2004 (<http://www.csm.ornl.gov/meetings/SCNEworkshop/Workshop-Report-ORNL-TM-2004-132.pdf>).

Contributions

Exemples de besoins en moyens de calcul pour des simulations « Grand Challenge » Contribution de G. Zerah (CEA/DAM)

Modélisation multi-échelles du comportement des métaux

L'objectif est de parvenir à une modélisation prédictive du comportement des solides, via l'utilisation de modèles décrivant plusieurs échelles spatiales et temporelles

La simulation multi-échelles des matériaux en mécanique vise à définir des modèles de comportement utilisables en calculs de type 'calcul de structures' (typiquement, un calcul d'ingénieur par éléments finis), à partir de données calculées à des échelles inférieures, depuis l'échelle atomique (10⁻¹⁰ m) jusqu'à l'échelle du cm. L'objectif est de parvenir à effectuer des prédictions de comportement dans des conditions où l'expérience n'est pas réalisable (grandes vitesses de déformation, grandes pressions, grandes variétés de sollicitations, typiquement).

On peut distinguer deux types de propriétés relatives au comportement : Les propriétés de type « équation d'état ». Elles relèvent de la thermodynamique à l'équilibre, et elles permettent de simuler les grandes déformations élastiques (c'est-à-dire réversibles) de divers matériaux, des métaux aux nanotubes de carbone. Les propriétés élastiques utilisées au niveau macroscopique par un code de type « éléments finis » peuvent être calculées « à la volée » au niveau microscopique avec chaînage direct des codes

Les propriétés liées à la présence de défauts. Ce sont les propriétés élastiques affectées par des impuretés ou bien les propriétés de déformation plastique (irréversibles) déterminées par les défauts linaires appelées « dislocations » qui balayent un cristal parfait lors de telles déformations. Ces défauts sont susceptibles d'un traitement à une échelle mésoscopique.

En général, le chaînage des échelles dans les codes s'effectue préférentiellement à travers l'introduction manuelle dans un code traitant l'échelle supérieure, de données calculées une fois pour toutes à l'échelle inférieure. Cependant, dans certains cas de localisation des défauts (rupture par fissures, notamment), on peut raffiner localement la description géométrique pour descendre de l'échelle nanométrique à l'échelle atomique (méthode de quasi-continuum).

Les défis actuels concernent le développement de modèles de comportement macroscopiques de polycristaux capables de prendre en compte la simulation de la déformation en conditions rapides (chocs), le caractère polycristallin des métaux (effets des joints de grains)

Etude de la réactivité chimique par une approche multi-échelles

Les modèles du futur.

La description de l'évolution de milieux réactifs sous l'effet de la pression ou de la température doit reposer sur une analyse fine des réactions chimiques. Cette étude nécessite des modèles de type cinétique chimique qui traitent de manière exacte la zone de réaction.

Possibilités actuelles.

L'outil de base présent et futur pour l'analyse à l'échelle microscopique des propriétés de la zone de réaction chimique est la dynamique moléculaire classique (compte tenu du nombre de molécules à prendre en compte pour cette application spécifique (de l'ordre du million), le calcul *ab initio* reste exclu). Cette méthode nécessite la connaissance du potentiel classique rendant compte des interactions entre les atomes du système. Les outils aujourd'hui disponibles permettent de simuler :

- des systèmes inertes par hypothèse comportant quelques millions de molécules,
- des systèmes réactifs simples de type AB comportant quelques millions de molécules,
- des systèmes réactifs complexes comportant quelques centaines de molécules.

Ces outils permettent la prédiction de quelques-unes des données nécessaires à la construction d'un modèle macroscopique de type cinétique (éléments de cinétique chimique) et apportent des informations qualitatives sur la physique de la zone de réactions (structure et propriétés thermodynamiques). Les ressources de calcul nécessaires sont au maximum d'une centaine d'heures sur quelques centaines de processeurs.

Les travaux futurs concerneront, dans une première phase de 3 à 5 ans :

1. Paramétrage de potentiels classiques : calculs *ab initio* standards sur quelques centaines de molécules.

2. Extraction de toutes les informations nécessaires à la construction de modèles de type cinétique *i.e.* calculs des propriétés thermodynamiques et de cinétiques chimiques avec une grande précision : quelques dizaines de milliers de molécules associées à un potentiel évolué, soit quelques dizaines de milliers d'heures de calcul sur un processeur actuel de la machine du CCRT.

3. Simulation d'un front de réaction chimique dans des géométries permettant de mettre à l'épreuve les modèles cinétiques : comparaison directe entre le calcul de dynamique moléculaire et le calcul macroscopique. Le calcul microscopique nécessitera quelques millions de molécules associées à un potentiel évolué soit quelques millions d'heures de calcul sur un processeur actuel.

Au-delà de 5 ans :

1. Calcul en routine de la propagation de réaction chimique dans un environnement complexe (interaction avec différents milieux connexes) avec un modèle de type cinétique.

3. Calcul loupe (défauts, points singuliers, etc.) directement par dynamique moléculaire : quelques centaines de millions de molécules associées à un potentiel évolué, soit quelques centaines de millions d'heures de calcul sur un processeur actuel.

Modélisation multi-échelle des effets d'irradiation

L'objectif est la connaissance de l'évolution des propriétés des alliages d'actinides en fonction du temps sous irradiation.

Les défauts élémentaires issus par exemple de la désintégration d'un actinide radioactif sont produits dans un temps inférieur à la nanoseconde, alors que leur diffusion et l'établissement d'un régime stationnaire se produisent sur des durées allant de la microseconde à la seconde, quant au temps « macroscopique » d'observation il doit s'étendre sur plusieurs décennies.

D'autre part, les défauts ponctuels produits occupent un volume de quelques mailles cristallines, les défauts étendus comme les dislocations concernent le grain alors qu'un changement de phase (de type martensitique) ou un phénomène de gonflement concernent l'échantillon dans sa globalité (le poly cristal).

Il est donc nécessaire d'adopter une démarche multi-échelle pour modéliser cette évolution de manière fiable. Cette démarche peut s'articuler autour de deux objectifs :

- mieux connaître les processus mis en jeu à chaque échelle de temps et d'espace ;
- alimenter en données les domaines les plus macroscopiques.

Les échelles les plus exigeantes sont les échelles atomiques et quantiques.

1- Etudes à l'échelle quantique

L'objectif des études à l'échelle quantique est de mieux connaître les processus élémentaires impliqués dans la stabilité des alliages d'actinides (en particulier le rôle des corrélations électroniques) et d'alimenter les échelles supérieures – en particulier l'échelle atomique- soit par le calcul de données thermodynamiques utiles pour ajuster un potentiel effectif, soit pour élaborer un modèle spécifique du potentiel effectif lui-même.

2- Etudes à l'échelle atomique par la dynamique moléculaire

Les études à l'échelle quantique sont limitées en taille à moins d'une centaine d'atomes. Pour espérer prédire les conséquences de l'irradiation sur les propriétés des alliages d'actinide il est nécessaire de prendre en compte, au minimum, un volume qui contient la totalité d'une cascade de déplacement induite par les atomes de recul issus de la désintégration des atomes d'actinide. Cette échelle ne peut être atteinte qu'en abandonnant la description quantique et en passant à une description atomique de type dynamique moléculaire. On attend de cette description la connaissance de la nature et de la quantité des défauts produits par auto irradiation.

Nanoélectronique
contribution de J. L. Leray, CEA/DIF/DCRE

Note : La contribution a été fournie à la mission sous forme de transparents qu'elle a résumés ci-après en raison de l'importance du sujet. Une proposition de programme incitatif fédérateur pluriannuel considérant à la fois la nanoélectronique et les matériaux a été transmise à la direction de la Recherche par le CEA. La suite qui pourrait être donnée relève de l'action de soutien « Calcul Intensif et Grilles » de l'Agence Nationale de la Recherche et du GIP qui doit la préfigurer.

Les réalisations matérielles relevant des Technologies de l'information et de la communication reposent sur les circuits sur silicium. Ceci peut représenter jusqu'à 30 % de la valeur d'un produit tel qu'un avion militaire ou civil ou qu'une automobile. Deux types de calculs sont requis : 1) matière (nanosciences, nanoprocédés = nanomatériaux « classique ») 2) information (courants d'électrons, de spins etc. = « transport »).

Roadmap » de l'ITRS (International Technology Roadmap of Semiconductors), tracée jusqu'en 2016, mais comportant des défis technologiques importants. L'historique récente indique que les années 1998-2000 ont assisté à une accélération du progrès par rapport à la loi empirique de Moore• On est déjà dans le nanométrique, en ce qui concerne les transistors (45-32 nanomètres en xy, 1-2 nanomètres en z).

La simulation des matériaux est un soutien essentiel aux activités d'assemblage des parties du transistor « front end » et « device » (fonctionnement électrique) principalement. Les défis sont : les clusters, l'évolution vers les transistors à faible nombre d'électrons, qui requièrent la parallélisation des calculs. Il faudra également étendre les codes *ab initio* actuels aux modèles de transport quantique et mixer les deux niveaux de calcul : *ab initio* parallélisé / transport électronique en parallèle.

Il faut maîtriser la variabilité pour assurer la fiabilité, et le rendement de la fabrication ce qui requiert des calculs en parallèle et supervisés. Ce type de considération est essentiel pour la réussite du transfert industriel de l'innovation dans ces domaines caractérisés par une grande réactivité dans la compétition et l'intensité capitalistique des investissements sur les nouvelles productions.

Au-delà de ce plan, il faut considérer des plans « révolutionnaires » qui se justifient par l'essoufflement du paradigme de la loi de Moore, le coût des usines de fabrication par rapport à la solvabilité des marchés. Plusieurs concepts de transition ou alternatifs sont à considérer : « nano-inside » et nano tubes de carbone, électronique moléculaire, spintronique. Dans ce cas, on devra affronter la difficulté à simuler l'état amorphe.

Tableaux synthétiques

Les tableaux synthétiques se réfèrent aux questions posées pour l'exercice de prospective :

- Préciser la contribution attendue de la simulation, du calcul intensif et de la manipulation de données massives dans votre domaine scientifique ou technologique ?
- Quel pourra être l'impact de la mise en œuvre des ressources décrites dans les scénarios ci-après ? D'autre part, il a été demandé aux contributeurs d'envisager les scénarios suivants :

	Scénarios			
Année	France A	France B	Europe C	Europe D
2006	25 teraflops	50 teraflops		
2009	90 teraflops	250 teraflops	0,5 petaflops	1 petaflops
2015	0,5 petaflops	2 petaflops	6 petaflops	12 petaflops
2020				100 petaflops

Pour permettre d'intégrer la réflexion sur l'utilisation éventuelle de machines d'architecture vectorielle, les scénarios suivants ont été ajoutés. Toutefois, le caractère spéculatif de l'analyse de l'offre de tels systèmes sur le marché a été précisé aux membres des groupes de réflexion prospective.

	Scénarios		
Année	France A	France B	Vectoriel E
2006	2 teraflops	5 teraflops	15 teraflops
2009	3 teraflops	10 teraflops	100 teraflops

Les scénarios ont été conçus dans la perspective suivante :

- Scénario A Evolution à budget constant en se contentant du progrès technologique
 Scénario B Rattrapage par la France suivant les recommandations de ce rapport
 Scénario C Programme européen, exécution minimale
 Scénario D Programme européen, exécution définie dans le projet tripartite

Problème/ Application	Challenge scientifique et apport scientifique	Scénario choisi / Année	Ressources requise (teraflops efficaces*heure) / Mode d'exploitation	Renvoi table 2
Fusion Contrôlée (CEA / DRFC)				
Ondes de chauffage dans ITER	Propagation et absorption d'ondes de chauffage additionnel Résolution spatiale fine. Calcul du chauffage dans ITER.	France B / 2010	1 000 teraflops*heure	
Instabilités dans les plasmas	Simulations d'instabilités MHD dans des plasmas en ignition. Résolutions spatiale et temporelle fines. Couplage aux particules rapides. Evolution non linéaire.	Europe C / 2010	100 000 teraflops*heure "Grand Challenge"	
Plasmas turbulents ITER	Transport turbulent dans les plasmas d'ITER. Turbulence développée. Description cinétique nécessaire : 5 D. Résolutions fines en espace, vitesse, et temps.	Europe C-D/2010	1 000 000 teraflops*heure	
Energie nucléaire (CEA /DEN)				
Cycle de vie des installations nucléaires	On résume ici la contribution CEA-DEN figurant ci-dessus. Les trois disciplines qui dominent pour les besoins en calcul intensif sont : les matériaux, la neutronique, la thermohydraulique.	France B Europe C ou D	Matériaux et thermohydraulique considérés « Grand Challenge »	

Climatologie / Institut Pierre Simon Laplace				
Climat futur de la terre	Quels sont les systèmes à prendre en compte pour déterminer le futur de la planète ? Modèle climat système terre avec toutes les composantes interactives permettant les phénomènes non linéaires.	France B Vectoriel parallèle 2006. Niveau de base : besoin dès 2005 Vectoriel Parallèle Scénario E 2009	En 2005 : 50 % de la machine à 5 teraflops (40 % d'efficacité). Permet 15 simulations de 40 ans en 3 mois sur 40 processeurs vectoriels en parallèle. Mémoire critique : 160 Go En 2009 : 25 % de la machine à 100 teraflops (40 % d'efficacité) 15 simulations de 200 ans en 3 mois sur 200 processeurs en parallèle. Mémoire critique : 200 Go	IPSL-8
Perturbations anthropiques et rétroactions nuageuses	Améliorer la connaissance de la réponse globale du climat aux perturbations anthropiques en améliorant les rétroactions nuageuses. Bonne représentation des nuages bas et de leur sensibilité aux changements de circulation, de température et d'humidité.	France B Vectoriel Parallèle 2006. Croissance CPU facteur 20 par rapport à ligne de base. Besoin dès 2005	Résolution verticale x2 à x4 couche limite. Résolution horizontale x4. CPU 50 % de la machine à 5 teraflops (40 % d'efficacité). 15 simulations de 200 ans en 3 mois sur 20 processeurs en parallèle chacune	IPSL - 1
Effet du couplage climat carbone à l'échelle globale.	Simuler les changements climatiques et les couplages climat-carbone à l'échelle globale. Idem ci-dessus avec en plus les modèles de carbone des surfaces continentales et de l'océan	France B Vectoriel Parallèle 2006. Croissance facteur CPU 30 et mémoire 2 par rapport à ligne de base. Besoin dès 2005	Résolution idem ci-dessus. CPU 50 % de la machine à 5 teraflops (40 % d'efficacité) 10 simulations de 200 ans en 3 mois sur 20 processeurs en parallèle chacune.	IPSL - 2
Quantifier les changements climatiques en Europe	Bien simuler les régimes dépressionnaires des moyennes latitudes. Effets des changements climatiques : tempêtes, inondations, sécheresse	France B Vectoriel Parallèle 2006. Croissance facteur CPU 15 et mémoire 10	Résolution horizontale 1°x1° (300 x 225) soit facteur 9. CPU 50 % de la machine à 5 teraflops (40 % d'efficacité) 10 simulations de 200 ans en 3	IPSL- 3

		par rapport à ligne de base. Besoin dès 2005	mois sur 20 processeurs en parallèle.	
Impact régional du changement climatique sur la chimie de l'atmosphère	Quantifier l'impact régional du changement climatique sur la chimie dans la troposphère et la stratosphère. Modèle couplé avec chimie-aérosols. Ajouter 200 traceurs et encore plus de réactions chimiques. Représentation onde de gravité. Prise en compte cirrus élevés et nuages polaires	France B Vectoriel Parallèle 2006 et Vectoriel Parallèle 2009. Croissance facteur CPU 100 et mémoire 40 par rapport à ligne de base.	Amélioration de la résolution verticale à la tropopause pour atteindre 100 niveaux Facteur 4 en résolution. CPU 50 % de la machine à 5 teraflops (40 % d'efficacité) 15 simulations de 40 ans en 3 mois sur 40 processeurs en parallèle Mémoire critique : 160 Go. En 2009, simulations plus longues.	IPSL - 4
Cyclones tropicaux	Simuler les phénomènes extrêmes dans les régions tropicales et leur évolution Simuler les cyclones tropicaux	France B Vectoriel Parallèle 2006 et Vectoriel Parallèle 2009. Croissance facteur CPU 90 et mémoire 40 par rapport à ligne de base.	Résolution horizontale 0,5°x 0,5° (150 x 110). CPU 50 % de la machine à 5 teraflops (40 % d'efficacité) 15 simulations de 40 ans en 3 mois sur 40 processeurs en parallèle En 2009, simulations plus longues ou à plus haute résolution	IPSL-5
Le devenir des phénomènes tropicaux	Le devenir des phénomènes tropicaux (El Niño, moussons, ...). Résoudre les transferts verticaux aux échelles pertinentes. Formulation explicite de certains phénomènes : déferlement ondes internes, turbulences de surface	Vectoriel Parallèle E 2009	Facteur 1 000 par rapport au couplé 1° atmosphère et 0,5° océan. Donc x 20 000 CPU 50 % de la machine à 100 teraflops (40 % efficacité) 15 simulations de 20 ans 100 processeurs en parallèle	IPSL-6

Risque de surprise climatique liée à la circulation thermohaline.	Quantifier le risque de surprise climatique liée à la circulation thermohaline. Modélisation des interactions océan-glace de mer-atmosphère à échelle de la dizaine de km, simulations longues, approche statistique	Vectorel Parallèle E 2009	Etape intermédiaire : 50 % de la machine à 5 teraflops (40 % d'efficacité) 8 simulations de 400 ans en 3 mois sur 40 processeurs en parallèle	IPSL-7
Effet de l'utilisation des sols et de l'urbanisation sur le climat.	Comment évaluer l'effet des changements d'utilisation des sols et de l'urbanisation sur l'évolution du climat ? Modèle climat système terre avec toutes les composantes interactives permettant les phénomènes non linéaires	Idem IPSL-8 ci-dessus	Idem IPSL- 8	IPSL-9
Savoir utiliser les prochaines générations de machines.	Nouveau modèle avec bases numériques différentes pour parallélisation sur milliers de processeurs. Basculer la production scientifique sur ce modèle lorsqu'il permettra les mêmes études que le précédent.	France B Scalaire Parallèle	Accès à un grand nombre de processeurs pour développement, avec temps de restitution court.	IPSL-10
Océanographie à haute résolution				
Variabilité de l'océan à l'échelle globale	Variabilité de l'océan à l'échelle globale : sensibilité aux forçages atmosphériques. Comment fonctionne la variabilité de l'océan telle qu'on l'observe aujourd'hui ? Modèle ORCA –LIM résolution 1/4° 46 niveaux	2005-2007	2 400 heures/an simulé sur 186 processeurs IBM (IDRIS) Besoin: 150 ans simulés/année civile, 2005-2007	OCEANS-1
Variabilité de l'océan à l'échelle régionale	Variabilité de l'océan à l'échelle régionale : physique des tourbillons et courants de bord, transports de polluants, interaction turbulente avec l'atmosphère Compréhension des interactions d'échelles fondamentales pour le climat, compréhension des mécanismes générateurs des courants océaniques. Modèle ORCA -LIM, résolution 1/4° et 46 niveaux + zooms dans régions clés	2006-2009	Coût triplé par rapport à ORCA025 seul. Utilisation de calculateurs parallèles scalaires ou vectoriels	OCEANS-2
Circulation océanique à l'échelle globale.	Circulation océanique à l'échelle globale : vers la résolution des échelles importantes pour lever les incertitudes liées aux paramétrisations Défi de la représentation explicite des tourbillons et de la formation des eaux profondes.	2008 et au-delà	Facteur ~ 40 en CPU par rapport à ORCA 1/4° et Facteur ~ 14 en Mémoire	OCEANS-1

	Mise au point d'un modèle global au 1/12° Modèle ORCA –LIM, résolution 1/12° et 46 niveaux			
Compréhension du rôle de l'océan comme puits et source de CO2	Modélisation directe (non paramétrisée) de l'action de la mésoéchelle océanique sur les cycles biologiques et la séquestration du carbone. Elaboration des modèles physiques et des modèles de biogéochimie	2005-2009	Le coût du modèle biogéochimique représente 1 à 5 fois le coût du modèle physique..	OCEANS-3
Meilleure prise en compte des mesures.	Ré-analyses globales Modèle opérationnel avec assimilation Intégrer les observations des dernières décades dans un système dynamique cohérent pour la compréhension du système climatique. Modèle ORCA-LIM , résolution 1/12° et 46 niveaux Assimilation de données Schéma SEEK	2008 et au-delà	Facteur ~ 80 en CPU et facteur ~ 20 en mémoire	OCEANS- 4
Méthodes multi résolution matériaux métaux et biologie				
Chimie fine et catalyse	Modélisation de complexes organométalliques pour des applications à la chimie fine et à la catalyse : effet des ligands et des paires d'ions Représenter des systèmes de grande taille, ioniques ou non en solution. Etudier la réactivité des différentes parties et leur influence réciproque.			MMR-Métaux
Catalyseurs greffés sur silice.	Modélisation de catalyseurs greffés sur surface de silice. Comparaison catalyse homogène/ catalyse supportée Difficulté de représenter un greffage non régulier sur une surface amorphe. Influence de la surface sur l'efficacité de catalyseurs.			MMR-Métaux

Batteries au lithium	Diffusion des ions lithiums dans les solides. Réorganisation du solide/application aux batteries au lithium Représentation de la restructuration du solide. Résilience du matériau. Propriétés électroniques			MMR-Métaux
Matériaux nucléaires de structure et de stockage	Couplage entre les briques de base ayant individuellement atteint la maturité en 2005 : <i>Ab initio</i> , Dynamique moléculaire, Monte Carlo. Exemples : 1)matériaux de structure : aide au choix des matériaux des centrales du futur (fission 4 ^{ème} génération). Matériaux de stockage (oxydes-verres)	France B complémentarité scalaire vectoriel appréciée		DEN-MMR
Modélisation prédictive du comportement des solides	Il s'agit de prédire le comportement macroscopique des solides à partir d'études multi-échelles partant des principes de base. Prise en compte des défauts et dislocations. L'intérêt est de traiter des cas qui ne se prêtent pas à l'expérimentation (rayonnement, longue durée, autres situations extrêmes)	France B Europe C	En 2006-2010 : 1 petaflops*heure par calcul, plus si anisotropie	DAM-MMR
Réactivité chimique	Description des l'évolution des milieux réactifs sous l'effet de la pression ou de la température à partir de l'analyse des réactions chimiques. Prise en compte d'environnement complexe. Construction de modèles cinétiques.	France B Europe C Europe D	5 petaflops*heure à 500 petaflops*heure	DAM-MMR
Modélisation des effets d'irradiation sur les alliages	Il s'agit de comprendre l'évolution des propriétés des alliages d'actinides dans le temps sous irradiation. Les phénomènes se déroulent sur des échelles de temps allant de la nanoseconde pour la désintégration d'un actinide radioactif, à la seconde pour l'établissement d'un régime stationnaire. La prédiction doit se faire sur plusieurs décennies.	France B Europe C	5 petaflops*heure par calcul en 2006-2010	DAM-MMR
Nano-électronique	Il s'agit de se donner les moyens de suivre l'évolution des nouveaux dispositifs électroniques, qui peut se faire de manière évolutive dans le cadre la « Roadmap ITRS » ou révolutionnaire. Dans les deux cas des défis importants de modélisation et de simulation se présentent.			CEA-Nanoélec

Senseurs et commutateurs biologiques, micro livraison de médicaments	Simuler la dynamique de l'auto-assemblage de nanostructures biologiques. Conception rationnelle d'outils nanotechnologiques tels que des senseurs et des commutateurs biologiques, des systèmes de micro-livraisons de médicaments.	France A		MMR - biologie
Mécanismes moléculaires de la biologie, des maladies. Conception de nouveaux médicaments	Simuler des systèmes biologiques : simuler la dynamique de l'assemblage et des transformations de macromolécules biologiques, et ensuite d'organelles et de cellules entières (organisation de la chromatine, repliement des protéines, assemblage des membranes, transduction du signal, formation des complexes). Comprendre la formation des complexes et la transmission d'information. Comprendre les mécanismes moléculaires des processus biologiques. Description au niveau moléculaire des maladies. Conception rationnelle de nouveaux médicaments.	France B		MMR - biologie
Enzymologie moléculaire.	Simulation de réactions enzymatiques. Prédiction métaboliques.	France B		MMR - biologie
Astrophysique : Projet HORIZON				
Galaxies	Physique du gaz à petite échelle, galaxies spirales et elliptiques, interactions, étoiles et vents	France B 2006 France B 2009 Europe D 2009	1000 / communauté 5000 / communauté 20000 / grand challenge	Horizon Galaxies

Amas de galaxies	Physique du gaz à grande échelle, amas de galaxies, structure interne et distribution à grande échelle	France B 2006 France B 2009 Europe D 2009	500/communauté 2500/communauté 10000/grand challenge	Horizon Amas de galaxies
Grandes structures	Formation des halos de matière noire et distribution à grande échelle, modèles semi-analytiques, surveys	France B 2006 France B 2009 Europe D 2009	250/communauté 1250/communauté 5000/grand challenge	Horizon Grandes structures
Astrophysique CEA				
Vision MHD 3D Soleil/ Etoiles	Comprendre la dynamique interne des étoiles et le Soleil archétype	2000-2005: CCRT/CEA 2005-2010: France B 2006 2010-2015: Europe D	2005-2010: 800 heures par run, machine scalaire (effectif 5 teraflops, dizaine de runs) 2010-2015: machine scalaire, 1500 heures/run (quelques dizaines de runs, résolution au moins 5 fois supérieures par dimension)	
Relation Soleil-Terre, interaction entre étoiles	Comprendre la variabilité du flux solaire et ses interactions avec la Terre	2005-2008 : calcul 1D/3D France B 2008-2012: observations / prédictions	Machine vectorielle Quelques runs de 20 heures (effectif 5 teraflops) Au-delà 2010 : Quelques centaines d'heures/run ?	
Combustion & Supernovae	Comprendre les mécanismes de fragmentation du milieu interstellaire jusqu'à la formation de cœurs denses préstellaires	France-B: 2006/2015Europe-D: 2009	2005-2010: Environ 30 teraflops*heure par simulation. 2010-2015: De 300 à 1 000 teraflops*heure selon la physique implémentée.	

Milieu Interstellaire	Comprendre les mécanismes de fragmentation du milieu interstellaire jusqu'à la formation de cœurs denses préstellaires	France-B: 2006/2015 Europe-D: 2009	2005-2010: Environ 30 teraflops*heure par simulation. 2010-2015 : De 300 à 1 000 teraflops*heure selon la physique implémentée.	
--------------------------	--	---------------------------------------	---	--

TABLEAU II : "Pré-requis, Blocages à lever"

Référence table 1	Pré-requis / Difficulté scientifique / point bloquant / besoin formation	Action à entreprendre	Planification souhaitable de l'action (Années début/fin)
Astrophysique : Projet HORIZON			
Horizon Galaxies	Une meilleure résolution implique une physique à petite échelle plus riche. Stockage : schéma d'indexation des données Scalabilité pour nombre CPU > 1000 Technique de parallélisation massive et de gestion des grosses bases de données	1- Développement de routines de physique (atomique, moléculaire, modèle sous-maille, diffusion des métaux, vents, trous noirs super-massifs) 2- Développement des logiciels d'analyse et de visu adaptés aux très gros runs 3- Tests des algorithmes de décomposition de domaine pour NCPU > 1000	1- 2005-2007 2- 2005-2007 3- 2005-2007
Horizon Amas de galaxies	idem	4- Développement de logiciels d'analyse en temps réels (cartes virtuelles)	4- 4- 2005-2007
Horizon Grandes structures	idem	5- Développement de modèles semi-analytique de post-traitement parallèles	5- 5- 2005-2007

Climatologie / Institut Pierre Simon Laplace			
IPSL-1 à 6	<ul style="list-style-type: none"> • Poursuite du développement de la physique des modèles • Augmentation de la résolution verticale et horizontale • Parallélisation du modèle sur 10^{aine} de processeurs • Banaliser les simulations d'ensembles • Formation au calcul parallèle de tous les étudiants • Besoin d'organisation pour préparer les • Organisation autour des ressources matérielles et humaines 	<ul style="list-style-type: none"> • Maintenir de bons temps de réponse pour les développements du modèle • Maintenir de bons temps de réponse pour les ajustements des paramétrisations lors de la réalisation des expériences de référence • Finaliser la parallélisation du couplé actuel • Mise en place outils pour simulations d'ensemble, travail main dans la main avec centres de calcul pour mise en place grosse machinerie incluant pré et post-traitements, soumissions enchaînées de jobs, ... • Maintenir ressources stockage et post-traitement • Faciliter exploitation résultats sur les différents sites de calcul 	<p>Au plus tôt Et à maintenir dans la durée</p>
IPSL 7-à 9	Faire des simulations climat avec le système complet impose des contraintes lourdes en CPU, mémoire, espaces fichiers, possibilités post-traitements...	Décision de s'équiper du calculateur permettant ce type d'étude à prendre maintenant	
IPSL- 10	<p>Equipe pluridisciplinaire</p> <ul style="list-style-type: none"> • Numéricien • Algorithmicien • Physicien • Informaticien • Chimiste 	<ul style="list-style-type: none"> • Compléter équipe actuelle pour avoir le « répondant » nécessaire aux sollicitations pour construire ce nouveau modèle • Détecter 4 ans à l'avance la disparition des calculateurs vectoriels pour fixer le rendez-vous avec le modèle climat suivant capable de performances sur des milliers de processeurs 	<ul style="list-style-type: none"> • A monter dès maintenant • Basculer production scientifique sur ce modèle lorsqu'il permettra des avancées scientifiques climat. • J-4ans
Océanographie à haute résolution			
OCEANS-1	<p>Prérequis : modèle océanique performant pour la représentation de la mésoéchelle et des interactions courant/topographie : OPA 9 est un outil performant à continuer à développer.</p> <p>Point bloquant : disposer des heures de calcul nécessaires</p>	<p>1 - optimiser l'accès aux ressources de l'IDRIS et aux autres calculateurs européens pour les gros projets</p> <p>2 - faire évoluer rapidement les moyens de calcul (scalaire et vectoriels) disponibles en France.</p>	<p>Immédiatement, développement à long terme : 2005-2012 et au-delà</p>

OCEANS-2	Pré-requis: outils pour la régionalisation (données aux frontières, logiciel tel que AGRIF (développé au LMC) permettant d'effectuer des zooms régionaux. Outils existants mais à développer fortement	Outre le problème de temps calcul, il y a un manque de personnel permanent (ingénieur) sur ces questions en France. Le travail est effectué trop souvent par du personnel temporaire (grande perte de temps en formation)	Immédiatement, développement à long terme : 2005-2012 et au-delà
OCEANS- 3	Prérequis : modèles bio-géochimiques performants	Renforcer pluridisciplinarité bio-géochimie / dynamique	Immédiatement, développement à long terme: 2005-2012 et au-delà
Méthodes multi résolution matériaux métaux et biologie			
MMR Métaux	Tous les processus de calculs mentionnés sont itératifs et la durée d'une itération doit être nettement inférieure à la limite de temps de chaque soumission. Pour des systèmes de grandes tailles, des limites ont été atteintes et le calcul devient inefficace.	Autorisation d'avoir des durées de calculs supérieures à 24h car l'augmentation du nombre de processeurs semble ne pas être toujours la solution efficace.	
MMR-biologie	Continuer le développement de modèles physiques. Continuer la parallélisation des logiciels Validation des modèles. Applications pour valider.		
DEN-MMR	Couplage entre briques de base Calculs de forces dans l'état excité Cascades de déplacement <i>ab initio</i> Conserver leadership européen dans les codes <i>ab initio</i>	Poursuivre investigation des méthodes de calcul de forces dans l'état excité Mettre en place une structure nationale d'accueil pour l'aide au développement. Formation en sciences « numériques » dans la filière « matériaux » Participation aux projets de logiciel libre (ABINIT, PWSCF, FPMD). Améliorer position dans les réseaux européens : Psi-k, CECAM (Lyon), ICTP (Trieste)	
DAM-MMR CEA- Nanoélec	Programme fédérateur « maîtrise du logiciel et du calcul intensif pour la simulation des matériaux et des nanostructures » soumis au Ministère de la Recherche et à l'ANR	- rendre plus performant les logiciels de base pour chacune des échelles, - faciliter l'intégration horizontale pour le calcul multi-échelles, -organiser des workshops et des formations spécialisées.	Action à lancer immédiatement, ne produira son effet que progressivement sur 5 ans, ce qui est l'échelle de temps la plus courte à considérer pour les développements de codes et de méthodes ambitieuses.
Astrophysique CEA			
Astro CEA	- Contraintes expériences Spatiales/ Problème des	-Renforcer assistance programmation	2000-2005: Mise en place équations, validation

I	instabilités numériques et physiques/ Topologie des champs magnétiques, adéquation des méthodes numériques à la complexité grandissante des problèmes / Formation aux techniques numériques, support	-Amélioration des techniques numériques pour résolution spatiale au-delà de 1000^3 et de la puissance de calculs pour simulations plus réalistes	code 2005-2010: Etudes des interactions convection-rotation- turbulence-magnétisme 2010-2015: calculs réalistes, physique plus riche (atmosphère, multifluides, conditions aux limites) 2015-2020: Evolution stellaire (sur des Gyr) d'étoiles 3D MHD, phases précoces et ultimes, liens dynamiques entre différentes phases, Machines scalaires + vectorielles ?
Astro CEA - II	Description atmosphère, flux solaire= $f(t, \dots)$ / transition champ B fort à faible/Formation aux techniques numériques, support	-Portabilité des codes locaux 1D enrichis sur machines vectorielles pour grands nombres de pas temporels -- Mise en place d'un protocole faisant le pont entre la dynamo solaire interne et le code magnéto-sphérique de la terre.	2005-2008: Mise en place du problème ; simulations prévisionnelles des observations 2008-2012 : extension au climat du passé : 1D, 3D ? 2012-2015: dépendra des précédentes étapes : non défini aujourd'hui
Astro CEA - III	-Étudier et maîtriser la micro-physique en œuvre dans les SNIa. -Réunir des experts des nombreuses disciplines impliquées (hydrodynamique, combustion, physique nucléaire, physique stellaire...) - Trouver des algorithmes parallèles performants pour modéliser cette physique sur des machines à plus de 1000 processeurs.	- Poursuivre et intensifier les collaborations transdisciplinaires, notamment à travers le projet « Supernovae et Combustion ». -Former des étudiants à ces disciplines et au numérique - Développer un support technique de haut niveau proche des physiciens.	2005-2010: Analyse spécifique des différents processus physiques et modélisation mono-dimensionnelle des supernovae 2007-2015: calculs multi-dimensionnels incluant au mieux les différents processus microphysiques.
Astro CEA-IV	L'objectif est de faire le lien entre les échelles galactiques et les cœurs denses de formation stellaire	- Poursuivre le développement d'un code spécifique et des outils d'analyse associés - Développer un support technique de haut niveau proche des physiciens	2005-2010: Simulation hydrodynamique avec de la gravité et de la chimie. 2007-2015: Simulation MHD. Prise en compte de manière réaliste de la turbulence.

Formation au calcul scientifique

<p>Note sur la formation en calcul scientifique contribution de Serge Petiton, CNRS /STIC</p>
--

Cette note synthétise certains points évoqués par le groupe de travail et propose quelques pistes possibles ; indépendamment de l'existant. Depuis de nombreuses années la question de la formation en calcul scientifique intensif est discutée régulièrement sans vraiment de réponses claires. A notre avis, le thème de la formation au calcul scientifique devrait faire l'objet d'une réflexion approfondie.

Les formations actuelles en calcul scientifique, intensif ou non, sont difficiles à lister et synthétiser. Elles sont rattachées à différents départements selon les universités ou écoles et n'apparaissent pas toujours en tant que telles dans les cursus. La mise en place récente des LMD ayant restructuré aussi ces enseignements, il est difficile de bien connaître et évaluer les nouveaux. De plus, les mêmes intitulés de cours et des contenus similaires peuvent correspondre à des formations très différentes selon les domaines de recherches des enseignants.

Ces formations doivent s'inscrire complètement dans le paysage du calcul scientifique en France. Si elles ne sont pas à la hauteur des ambitions affichées par ailleurs, l'efficacité des efforts faits sera moindre.

Formation en Licence et première année de Master

Constatant que de nombreux étudiants des filières universitaires scientifiques ne connaissent que très peu les bases en calcul scientifique en sortie des masters, le plus important semble être d'enseigner un « **socle commun de connaissance en calcul scientifique** » aux étudiants de Licence ou de première année de Master (dans le cadre de la structure LMD). On peut penser qu'il n'est pas nécessaire de faire ces cours en Licence mais uniquement en première année de Master. Néanmoins, les connaissances scientifiques nécessaires à l'acquisition de ce socle commun sont suffisantes dès la Licence pour presque l'ensemble des étudiants des disciplines scientifiques. Les élèves-ingénieurs ayant généralement une formation en calcul numérique et en informatique plus complète sont souvent mieux préparés pour le calcul scientifique. Le contenu de ce socle commun est à préciser mais devrait inclure des notions de méthodes numériques, d'informatique scientifique (au moins sur l'arithmétique flottante et les fonctions élémentaires), de génie logiciel et de formation à quelques outils classiques liés au calcul scientifique. Un nombre d'heure suffisant de travaux pratiques doit être proposé car ce domaine ne peut être uniquement enseigné magistralement.

Formation en seconde année de Master

Sur ce socle commun de connaissance en calcul scientifique doit ensuite reposer la formation en calcul scientifique de seconde année de Master (*professionnel* ou *recherche*). Celle-ci doit prendre des formes variées selon la discipline. Des masters à vocations purement « calcul scientifique » ou « calcul haute performance » doivent se développer, mais l'ensemble des étudiants en master scientifique doit pouvoir suivre des cours sur ces deux thèmes au sein des écoles doctorales ; dans des enseignements transversaux ou optionnels, par exemple. Ces enseignements doivent être pluridisciplinaires et non concerner que certains aspects du calcul scientifique. Des expérimentations sur machines adéquates doivent être proposées ; au niveau des mésocentres ou des centres de calcul des universités et écoles. Ces formations concernent principalement les masters *recherches* mais il est aussi nécessaire de développer des enseignements adéquats dans les masters *professionnels* pour former les ingénieurs en « génie calcul scientifique » nécessaires dans ce domaine.

Formation doctorale et post-doctorale

Les formations proposées doivent être aussi accessibles aux doctorants. Des enseignements plus « techniques », comme ceux enseignés à présent par les centres de calculs nationaux, doivent être proposés aux doctorants qui doivent utiliser des machines hautes performances. Un label « calcul scientifique » ou « calcul haute performance » peut être donné à certaines thèses de doctorat. Ce type de label a été proposé lors de discussions au sein d'ORAP (mais sans recommandations particulières pour l'instant). Ce label pourrait, par exemple, être demandé pour certains types de financement post-doctoral qui serait mis en place pour favoriser des recherches pluridisciplinaires.



Prospective sur le calcul intensif : la vision de l'INRIA

21 février 2005

STÉPHANE LANTERI, Directeur de Recherche INRIA
Responsable permanent du projet CAIMAN (UR Sophia Antipolis)

Table des matières

1	Introduction	1
2	L'INRIA et le calcul intensif aujourd'hui	3
3	La stratégie de l'INRIA vis-à-vis des moyens de calcul intensif	5
3.1	Les moyens de calcul des centres nationaux CINES et IDRIS	5
3.2	Les mésocentres	5
3.3	Les clusters de PC	5
3.4	Les grilles informatiques	6
4	La prospective sur l'accompagnement par la recherche	7
5	Synthèse	8

Préambule. Ce texte a été rédigé en collaboration avec un ensemble de responsables scientifiques et membres de projets de l'INRIA dont les activités ont à voir avec le calcul intensif¹ et plus particulièrement, avec la résolution sur plate-formes de calcul parallèles et distribuées de problèmes issus du calcul scientifique².

1 Introduction

Nous débutons cette analyse par un bref rappel historique concernant l'implication de l'INRIA dans le domaine du calcul intensif. Depuis le début des années 80, l'INRIA a toujours joué un rôle de précurseur dans ce domaine. L'INRIA, soutenait notamment la création de projets comme CAPRAN (calcul parallèle numérique), CALCPAR (calcul parallèle) et API (architectures parallèles intégrées) dont les activités étaient centrées sur les architectures et l'algorithmique des calculateurs parallèles. Néanmoins, à cette époque, le calcul vectoriel représentait la seule

¹Par *calcul intensif*, on entend ici la résolution sur ordinateur de problèmes nécessitant de grandes capacités de traitement (puissance de calcul, capacité mémoire et capacité de stockage sur supports fixes). Suivant les architectures de plate-forme, le calcul intensif fait appel au *calcul vectoriel* (en voie de disparition), au *calcul parallèle* qui repose sur l'association de plusieurs entités de calcul autour d'un réseau local (LAN - Local Area Network/SAN - System Area Network) et au *calcul distribué* qui repose sur l'association de plusieurs entités de calcul, éventuellement hétérogènes, distribuées géographiquement (réseau WAN - Wide Area Network).

²Dans ce document, le terme *calcul scientifique* fait référence à l'utilisation d'un ordinateur en tant qu'outil de travail dans une discipline scientifique quelconque. Il sert une démarche scientifique visant à l'élaboration d'une théorie ou à la confrontation avec la réalité. Il sert parfois à retrouver numériquement des résultats d'expériences. Une activité de calcul scientifique type combine des aspects de modélisation physique et mathématique, d'analyse numérique, de conception et dévaluation d'algorithmes et de programmation informatique, tout ceci dans le contexte de la résolution d'une problématique scientifique donnée.

alternative viable pour la résolution des grands défis posés par le calcul scientifique. Les industriels de l'informatique qui proposaient des solutions matérielles aux utilisateurs concernés étaient essentiellement l'Américain Cray et les Japonais Fujitsu, Nec et Hitachi. En France, la principale ressource de calcul vectoriel était fournie et administrée par le CCVR. Plusieurs projets de calcul scientifique de l'INRIA comptaient parmi les principaux utilisateurs des machines Cray du CCVR. Dans le même temps, l'INRIA s'intéressait de près aux nouvelles architectures de calcul parallèle comme une alternative aux calculateurs vectoriels. Plusieurs sociétés Américaines proposaient des machines de ce type et un petit nombre de systèmes furent installés en France, pour la majeure partie à l'INRIA. Il s'agissait de la CM-2/CM-200 de la société Thinking Machines Corporation (à l'INRIA Sophia Antipolis, machine SIMD comprenant 65536 processeurs 1 bit), la KSR-1/KSR-2 de la société Kendall Square Research (à l'INRIA Rocquencourt, machine MIMD disposant d'un système de mémoire virtuelle partagée tout à fait innovant) et les iPSC et Paragon produites par la division Super Computing d'Intel (à l'IRISA, machines à mémoire distribuée basées sur le processeur Intel i860). La grande diversité de ces architectures et le manque de standardisation des environnements logiciels et outils de programmation parallèle sont les deux causes principales dans la disparition prématurée de ces acteurs (et d'autres encore qui ne sont pas cités ici pour faire court) de l'industrie du calcul haute-performance. Par exemple, le concept SIMD à la base des calculateurs parallèles mis au point par la société TMC était dans les faits bien adapté à une classe relativement limitée de problèmes dont la majeure partie des applications du calcul scientifique ne faisait pas partie. De même KSR était en avance sur son temps avec une solution matérielle ad-hoc. Enfin, au milieu des années 90, la division Super Computing d'Intel disparaissait avec la montée en puissance des PC et l'apparition des premières grappes de PC.

Par la suite (milieu des années 90), le paysage du calcul haute-performance a largement bénéficié de plusieurs initiatives de standardisation tant au niveau de la programmation parallèle (l'exemple le plus frappant est la définition du standard de programmation par échange de messages MPI) que de celui des architectures des calculateurs (abandon du SIMD au profit du MIMD, conception de systèmes basés sur des processeurs non propriétaires, etc.). Des constructeurs comme Cray, SGI et IBM ont alors proposé des architectures qui ont vraiment stabilisé le domaine du calcul intensif. Cependant, ces super-calculateurs parallèles restent des systèmes onéreux à l'acquisition comme à la maintenance. Les centres nationaux IDRIS (successeur du CCVR) et CINES ont permis aux projets de calcul scientifique de l'INRIA de maintenir un haut niveau de réalisations dans leurs domaines d'activité respectifs. Néanmoins, certaines de ces activités et celles de projets de l'INRIA travaillant dans d'autres domaines pas forcément liés au calcul scientifique, supposent des conditions d'accès aux ressources qui ne sont généralement pas compatibles avec les règles de bon fonctionnement des centres nationaux. Les mésocentres sont apparus comme une alternative possible d'autant plus que l'INRIA est devenue partenaire du Centre Charles Hermite équipé d'une machine SGI Power Challenge Array (plus tard remplacée par une SGI Origin 2000 suivie d'une Origin 3800).

Parallèlement à l'accès aux super-calculateurs des centres nationaux et des mésocentres³, l'INRIA a largement investi (en moyens et en activités de recherche) dans une approche architecturale du calcul intensif (parallèle) complètement différente et prometteuse : les clusters (ou fermes, grappes) de PC. Ces plate-formes de calcul parallèles sont bâties autour de processeurs standards fabriqués en grandes séries (Intel Pentium et Itanium, AMD Athlon et Opteron) et de réseaux rapides (Gigabit Ethernet, Myrinet, SCI, etc.). Plusieurs plate-formes de ce type ont été installées à l'INRIA avec des configurations allant de quelques dizaines à quelques centaines de processeurs. Là encore, l'institut a joué un rôle moteur et de nombreux projets ont produit des résultats qui ont grandement contribué à promouvoir l'adoption de ce type de plate-forme (y compris dans le milieu industriel).

Globalement, ces différentes expériences ont permis à l'INRIA d'acquérir un haut niveau d'expertise dans différentes thématiques qui couvrent tout le spectre du calcul intensif, de l'informatique parallèle et distribuée, depuis les couches logicielles les plus basses (systèmes d'exploitation, protocoles réseaux, exécutifs pour les langages parallèles, gestionnaires de ressources, etc.) jusqu'aux applications (du calcul scientifique ou d'autres domaines) en passant par les langages de programmation parallèle, l'algorithmique numérique parallèle, les intergiciels pour l'aide au développement, à l'optimisation et à l'exploitation d'applications de calcul intensif, etc. De plus, l'INRIA a toujours veillé à porter son attention sur des architectures de calcul parallèles et distribuées innovantes. L'illus-

³Auxquels s'ajoutent des accès à d'autres ressources en France (au CEA notamment) et aux Etats-Unis dans le cadre de collaborations entre équipes de différents organismes.

tration la plus récente est sans aucun doute la participation active de l'institut au programme national GRID'5000 sur les grilles informatiques.

Dans la suite de ce document, on présente dans un premier temps la situation actuelle quant aux activités de l'INRIA en liaison avec le calcul intensif (section 2). Dans cette section, on s'est volontairement concentré sur les développements et applications traditionnelles du calcul scientifique. D'autres disciplines représentées à l'INRIA et utilisatrices de moyens de calcul intensif sont néanmoins mentionnées car elles sont susceptibles de requérir dans un avenir proche des capacités de traitement de tout premier ordre. Ces disciplines émergentes du calcul intensif sont de nouveau discutées dans la section 3 où l'on précise la stratégie de l'institut vis-à-vis des moyens de calcul intensif et plus généralement, des plate-formes de calcul parallèles et distribuées. La section 4 constitue la contribution de l'INRIA à la prospective sur l'accompagnement par la recherche dans l'utilisation des moyens de calcul intensif. Enfin, la section 5 résume cette analyse et en particulier les positions et propositions de l'INRIA vis-à-vis de la prospective sur le calcul intensif.

2 L'INRIA et le calcul intensif aujourd'hui

Plusieurs projets de l'INRIA sont concernés par le calcul intensif. Il s'agit en premier lieu des projets du thème **Num** (Systèmes Numériques) et en particulier, des projets des sous-thèmes **Num B** (grilles et calcul haute-performance) et **Num D** (modélisation, simulation et analyse numérique). Les projets concernés sont :

Num B : GRAND-LARGE (calcul parallèle et distribué à grande échelle), GRAAL (algorithmique et ordonnancement pour plates-formes hétérogènes distribuées), PARIS (programmation des systèmes parallèles et distribués pour la simulation numérique à grande échelle), SAGE (simulations et algorithmes sur des grilles de calcul appliqués à l'environnement), SCALAPPLIX (schémas et algorithmes hautes performances pour les applications scientifiques complexes).

Num D : CAIMAN (calcul scientifique, modélisation et analyse numérique), CALVI (calcul scientifique et visualisation), ESTIME (estimation de paramètres et modélisation en milieu hétérogène), MACS (modélisation, analyse et contrôle pour le calcul des structures), MICMAC (méthodes et ingénierie du calcul multi-échelle de l'atome au continuum), ONDES (modélisation et simulation de phénomènes de propagation d'ondes), OPALE (optimisation et contrôle, algorithmiques numériques et intégration de systèmes complexes multi-disciplinaires régis par des EDP⁴), SMASH (simulation, modélisation, analyse de systèmes hétérogènes).

Le projet IDOPT (identification et optimisation de systèmes physiques) du sous-thème **Num C** (modélisation et résolution de problèmes inverses en stochastique ou en grande dimension) est aussi un utilisateur du calcul intensif pour des applications du calcul scientifique. Les problématiques scientifiques qui intéressent ce projet (mais aussi le projet OPALE) nécessitent la réalisation de plusieurs simulations qui concourent toutes à la résolution d'un même problème. Il peut s'agir d'un problème d'optimisation (par exemple, la recherche de la forme optimale d'une aile d'avion en fonction de critères portant sur ses performances aérodynamiques) ou la résolution d'un problème inverse à des fins d'identification de valeurs de paramètres caractérisant le phénomène physique sous-jacent. On conçoit aisément que la résolution de ces problèmes peut requérir des durées de simulation très conséquentes dès lors que l'on traite de problèmes définis en 3 dimensions d'espace. Il est alors opportun de se demander si la mobilisation induite de ressources de calcul, nécessairement haute-performance, est compatible avec les règles de bon fonctionnement de centres nationaux comme l'IDRIS et le CINES.

Il convient par ailleurs de noter que plusieurs projets de l'INRIA (en nombre régulièrement croissant), associés à d'autres thèmes, font eux aussi appel au calcul intensif. Parmi eux, nous citerons un certain nombre de projets du thème **Bio** (systèmes biologiques) : ANUBIS (outils de l'automatique pour le calcul scientifique, modèles et méthodes en bio-mathématique), EPIDAURE (imagerie et robotique médicale), ODYSSEE (vision algorithmique et biologique), REO (simulation numérique d'écoulements biologiques) et SYMBIOSE (systèmes et modèles biologiques, bio-informatique et séquences). Mentionnons notamment qu'une des activités du projet SYMBIOSE concerne la parallélisation des traitements coûteux en génomique pour en accélérer fortement l'exécution. La mise

⁴Equations aux Dérivées Partielles.

en oeuvre des algorithmes développés dans ce projet vise les super-calculateurs parallèles mais aussi les grilles de calcul voire les architectures spécialisées. Le projet EPIDAURE fait appel au calcul parallèle et distribué dans la mise au point d'algorithmes performants pour la comparaison d'images médicales à des fins, par exemple, de planification d'interventions en neurochirurgie. Le projet ANUBIS a des activités de modélisation en dynamique des populations pour lesquelles des outils de simulation parallèles sont développés en étroite collaboration avec le projet SCALAPPLIX. Les activités du projet ODYSSEE touchent au domaine des neurosciences. Ce projet s'intéresse notamment à des questions liées à la modélisation de l'activité corticale. L'ensemble de ces projets représentent des disciplines émergentes du calcul intensif et deviendront à coup sûr, dans un avenir proche, des utilisateurs de tout premier ordre du calcul intensif.

Suivant les projets, les activités en relation avec le calcul intensif portent sur :

- la simulation numérique parallèle et distribuée pour traiter d'applications dimensionnantes (ceci recouvre des aspects de modélisation physique et mathématique, mise au point et analyse de méthodes numériques, programmation parallèle et distribuée, visualisation scientifique),
- l'algorithmique haute-performance (algèbre linéaire parallèle pour la résolution de systèmes denses ou creux, décomposition de domaine, méthodes multi-grilles parallèles, couplage de modèles, couplage de méthodes numériques),
- les intergiciels pour le calcul haute-performance et les grilles de calcul (outils et langages de programmation pour le calcul parallèle et distribué, exécutifs pour les langages parallèles, techniques et outils pour la distribution/régulation des calculs sur plate-formes hétérogènes).

Les principaux domaines scientifiques dont relèvent les applications dimensionnantes considérées par les projets des thèmes Num B et Num D mentionnés plus haut sont : la mécanique des fluides (aérodynamique, écoulements complexes multi-fluides et multi-phasiques) et des biofluides (écoulements dans le système cardiovasculaire), l'électromagnétisme (furtivité, compatibilité électromagnétique, bio-électromagnétisme), l'acoustique, l'aéroacoustique (propagation du bruit dans un écoulement), la géophysique (terrestre et sous-marine, endommagement des structures), la mécanique des structures (stabilité des structures), l'hydrogéologie (transport de contaminants dans le sous-sol, déplacement d'hydrocarbures), la physique des plasmas et des faisceaux de particules (fusion thermo-nucléaire, nanophysique), la chimie quantique moléculaire, la dynamique moléculaire.

Ces applications concernent essentiellement la simulation numérique de phénomènes physiques complexes, nécessitant la manipulation de grands volumes de données. Il s'agit de plus en plus souvent de considérer des problèmes en 3 dimensions d'espace (ou plus dans le cas du couplage EDP/faisceaux de particules) portant sur des géométries réalistes (dans certains cas, d'intérêt industriel) et des milieux hétérogènes. La multi-disciplinarité est une caractéristique quasi-récurrente des études sous-jacentes dont la mise en oeuvre est souvent facilitée par la proximité, au sein de l'INRIA, de différentes communautés. Il s'agit par exemple de collaborations entre :

- numériciens spécialistes de la mécanique des fluides et de la de la mécanique des structures, spécialistes du traitement d'images médicales et spécialistes de la modélisation géométrique pour diverses applications touchant au biomédical (modélisations numériques de l'activité électro-mécanique du coeur, de l'activité corticale, de l'écoulement du sang dans le système cardio-vasculaire) et au biohasard (modélisation numérique de l'interaction entre un champ électromagnétique issu d'un téléphone mobile et les tissus de la tête de l'utilisateur).
- numériciens spécialistes de la mécanique des fluides des milieux poreux et informaticiens spécialistes du calcul distribué pour des applications liées à l'environnement, en hydrologie.
- physiciens spécialistes des plasmas, numériciens spécialistes du couplage entre champs électromagnétiques et faisceaux de particules et informaticiens spécialistes du calcul distribué pour des applications touchant à la fusion thermo-nucléaire.

Il est instructif de noter que la majeure partie des ces collaborations multi-disciplinaires ont été initiées dans le cadre d'actions spécifiquement destinées à les promouvoir : ARC (Actions de Recherche Coopératives de l'INRIA) et ACI (Actions Concertées Incitatives) GRID du Ministère de la Recherche. Notons enfin qu'outre des projets INRIA, certaines des ces études impliquent aussi des équipes extérieures à l'INRIA (du CNRS et de l'INSERM notamment, suivant les études).

3 La stratégie de l'INRIA vis-à-vis des moyens de calcul intensif

Les projets de l'INRIA dont il est question à la section 2 font actuellement appel à différents moyens de calcul intensif que l'on peut décliner comme suit :

- les super-calculateurs parallèles à mémoire distribuée ou partagée des centres nationaux CINES et IDRIS,
- les machines parallèles de certains mésocentres (Centre Charles Hermite à Nancy, Université Bordeaux 1, Pôle de Calcul Intensif de l'Ouest),
- des clusters de PC des différentes UR (INRIA Rhône Alpes, INRIA Sophia Antipolis, IRISA, INRIA Lorraine, INRIA Futurs).

L'utilisation de ces différentes plate-formes de calcul est précisée dans les 3 paragraphes qui suivent. Dans le dernier paragraphe de cette section, nous proposons une vision prospective sur ce que devrait être la plate-forme de calcul idéal pour traiter les applications de calcul intensif du futur.

3.1 Les moyens de calcul des centres nationaux CINES et IDRIS

Il y a deux raisons principales qui motivent l'utilisation des moyens de calcul des centres nationaux CINES et IDRIS par des projets de l'INRIA :

- la réalisation de simulations numériques sur des cas de calcul réalistes afin d'obtenir des résultats de référence pour la problématique scientifique sous-jacente,
- l'évaluation des performances parallèles d'algorithmes haute-performance.

Les demandes de ressources correspondantes portent sur quelques milliers d'heures CPU ; les projets SCALAPPLIX (16000 heures en 2003 et 2004, 30000 heures demandées pour 2005) et SMASH (16000 heures en 2004 et 30000 heures demandées pour 2005) sont actuellement le plus gros consommateurs. Les utilisateurs de l'INRIA apprécient la stabilité des supercalculateurs gérés par les centres nationaux, la qualité de l'administration et du support assurés par le personnel de ces centres et, bien-sûr, les capacités de traitement offertes par ces systèmes. Il n'est cependant pas toujours aisé d'accéder à des configurations d'exécution de l'ordre de la centaine de processeurs compte tenu du grand nombre d'utilisateurs de ces systèmes et des politiques de gestion des ressources adoptées par les centres nationaux.

3.2 Les mésocentres

La difficulté soulevée en fin du paragraphe précédent n'est en général pas rencontrée sur les machines des mésocentres dont la haute disponibilité est une des raisons qui motivent leur utilisation. Ces systèmes, aux configurations plus modestes (de quelques dizaines de processeurs), sont alors utilisés pour valider les développements et obtenir des résultats préliminaires sur des cas de calcul intermédiaires.

3.3 Les clusters de PC

Les plate-formes de calcul les plus utilisées par les projets de l'INRIA consistent en des clusters de PC ou de systèmes multi-processeurs à base de processeurs standards (Intel Pentium ou Itanium, AMD Opteron) disponibles dans presque toutes les UR. Il y a plusieurs raisons qui expliquent cette situation :

- tout d'abord, ce type de plate-forme rassemble une bien plus large communauté de chercheurs que simplement ceux appartenant aux projets des thèmes Num B et Num D. Des projets impliqués dans des thématiques autres que celles qui relèvent du calcul scientifique sous sa forme la plus traditionnelle⁵ ont déjà été mentionnés à la section 2. Ces projets, comme ceux du thème **Bio** par exemple, visent des applications du calcul intensif pour lesquelles il est nécessaire de traiter de larges volumes de données à l'aide

⁵Rappelons ici les intitulés des comités thématiques des centres nationaux IDRIS et CINES : 1. environnement, 2. mécanique des fluides, 3. milieux réactifs, 4. astrophysique, géophysique et terre solide, 5. électromagnétisme et plasmas chauds, 6. mathématiques, mathématiques appliquées et systèmes mobiles, 7. systèmes moléculaires organisés et biologie, 8. chimie quantique et modélisation moléculaire, 9. physique, chimie et propriétés des matériaux.

d'algorithmes (pas forcément numériques) parallèles et distribués. D'autres projets travaillent sur des aspects plus bas niveau de l'utilisation de ce type de plate-forme (systèmes d'exploitation, protocoles réseaux, exécutifs pour les langages parallèles, gestionnaires de ressources, etc.) ou sur conception d'intergiciels pour l'aide au développement, à l'optimisation et à l'exploitation d'applications de calcul intensif. D'autres enfin se concentrent sur la mise au point, l'analyse et l'évaluation d'algorithmes numériques parallèles pour la résolution de certains problèmes génériques du calcul scientifique (résolution de systèmes linéaires denses ou creux, etc.).

- le transfert vers d'autres communautés d'utilisateurs, du milieu académique ou du privé, est une composante importante de valorisation des travaux réalisés par les projets de l'INRIA. Si l'on se limite aux activités des projets mentionnés à la section 2, il est instructif de noter que peu d'acteurs du privé (grands organismes, PME/PMI, etc.) exploitent des supercalculateurs parallèles intégrés tels que ceux des centres nationaux. Au contraire, les plate-formes de type ferme de PC (en particulier les clusters de machines SMP) sont souvent adoptées en remplacement de systèmes multi-processeurs classiques, souvent plus onéreux. Cette tendance résulte, d'une part, de l'apparition dans le paysage du calcul haute-performance de sociétés spécialisées, de plus en plus nombreuses, proposant des fermes de PC clés en mains (noeuds de calcul, réseaux rapides, systèmes d'administration et de gestion de ressources, compilateurs, etc.) et, d'autre part, de la disponibilité (sous forme libre ou payante) de nombreux outils et environnements logiciels dédiés au calcul intensif⁶, résultant de travaux de recherche dont certains ont été ou sont actuellement menés à l'INRIA ou impliquant des projets de l'INRIA⁷.

A l'avenir, l'INRIA continuera à utiliser ces trois types de plate-forme de calcul qui répondent clairement à des besoins différents. En particulier, l'institut reste très concerné par la question d'avoir accès dans un avenir proche à des capacités de calcul et de manipulation de données téraflopiques tels que celles que pourraient offrir les centres nationaux CINES et IDRIS.

3.4 Les grilles informatiques

S'il est clair qu'il est indispensable de préserver l'accès à des moyens de calcul et des services comme ceux offerts par les centres nationaux CINES et IDRIS, nous n'en sommes pas moins convaincus que ceux-ci ne couvrent qu'une partie des besoins actuels et futurs en calcul intensif. Il nous semble notamment que les ressources informatiques de ces centres et les modes d'utilisation de ces ressources ne sont pas compatibles ou ne fournissent pas de réponse satisfaisante aux problématiques suivantes (pour nommer les principales) :

- l'interactivité : le mode d'utilisation des moyens de calcul des centres nationaux ne permet pas une interaction avec le calcul comme par exemple, le couplage quasi-temps réel entre la simulation et la visualisation des résultats. Or, un tel couplage a une utilité évidente lorsqu'il s'agit de simuler et visualiser des phénomènes physiques fortement instationnaires en évitant de stocker sur disques des volumes conséquents de données correspondant à des instantanés du calcul.
- la disponibilité : les applications faisant appel à un processus d'optimisation (calcul optimal de formes) ou nécessitant la résolution de problèmes inverses (identification de paramètres) impliquent de mobiliser les ressources de calcul sur des durées conséquentes qui peuvent facilement dépasser les limites garantissant un bon fonctionnement des centres nationaux (au sens d'une exploitation équitable des ressources par tous les utilisateurs accueillis).
- le stockage : bon nombre d'applications traditionnelles (physiques des particules) ou émergentes du calcul intensif (traitement d'images pour des applications médicales ou vidéo) nécessitent de manipuler de très grands volumes de données. Pour ces applications, il est préférable de faire appel à des serveurs spécialisés combinant capacité de stockage téraoctique (au minimum) et accès rapides aux données.

⁶Par exemple, il existe aujourd'hui plusieurs environnements logiciels ou bibliothèques dédiés à la résolution numérique parallèle d'EDP et disponibles dans le domaine publique. PETSc (Portable Extensible Toolkit for Scientific Computing) développé au Argonne National Laboratory est l'exemple le plus remarquable.

⁷Comme MUMPS (a MULTifrontal Massively Parallel sparse direct SolverSL) et PaStiX (Parallel Sparse matrix package).

En somme, la plate-forme de calcul idéale devrait combiner des ressources de calcul haute-performance, de visualisation et de stockage. En toute généralité, ces ressources n'ont pas vocation à être localisées sur le même site géographique. Idéalement, ces ressources sont distribuées sur plusieurs sites géographiques interconnectés par des réseaux WAN (Wide Area Network). Par ces quelques lignes, nous venons de décrire les ingrédients et caractéristiques d'une grille informatique. Plusieurs projets de l'INRIA, appartenant aux thèmes **Num** (systèmes numériques), **Com** (systèmes communicants), plus précisément du sous-thème **Com A** (systèmes distribués et architectures réparties) et **Bio** (systèmes biologiques), ont des activités qui visent l'utilisation de grilles de calcul. L'exploitation efficace de ce type de plate-forme soulève de nombreuses questions liées à l'hétérogénéité des ressources (noeuds de calcul et réseaux d'interconnexion), la multi-localisation, la volatilité des ressources, la sécurité des données, l'algorithmique (numérique ou non) parallèle et distribuée, etc. auxquelles s'attaquent ces projets. La grille informatique au coeur du programme national GRID'5000, initié et soutenu par le Ministère de la Recherche, constitue un instrument quasi-unique en son genre, support essentiel aux activités de recherche des projets concernés et qui préfigure peut-être ce que sera la plate-forme de calcul massivement parallèle du futur. Même si GRID'5000 est une grille de calcul expérimentale et en aucun cas une plate-forme de calcul intensif, elle reste éventuellement reconfigurable pour permettre, par exemple, l'évaluation de performances d'algorithmes numériques parallèles et distribués. De ce point de vue, GRID'5000 est aussi une plate-forme qui permet de valider et évaluer en grandeur nature les algorithmes et outils du calcul scientifique intensif du futur.

4 La prospective sur l'accompagnement par la recherche

Une des questions à l'origine de la mise en place du groupe de travail coordonné par messieurs Héon et Sartorius concerne la stratégie à mettre en oeuvre pour doter la recherche française de moyens de calcul intensif comparables à ceux des grandes nations européennes et éventuellement capables de concurrencer (dans le cadre d'une structure alliant au minimum l'Allemagne, la France et le Royaume Uni) les configurations téraflopiques des Etats-Unis et du Japon. Il est clair que cette mise à niveau des moyens de calcul devra s'accompagner de réalisations logicielles et applicatives permettant de tirer les meilleurs profits de ces nouvelles capacités de traitement. Bien que dotée de nombreuses équipes très actives dans le domaine du calcul intensif, notamment pour ce qui concerne la réalisation d'outils et d'environnements logiciels pour le calcul parallèle et distribué ou encore la mise au point d'algorithmes numériques adaptés aux plate-formes de calcul intensif, la recherche française souffre d'un certain retard dans la résolution des grands défis scientifiques qui nécessitent des capacités de traitement téraflopiques voire pétaflopiques. La position de l'INRIA sur cette question peut se décliner en plusieurs remarques et propositions :

- notre expérience dans ce domaine montre une demande croissante du monde académique (et du monde industriel) pour des outils et environnements logiciels permettant la résolution de problèmes de plus en plus gourmands en puissance de calcul et en capacité mémoire. A titre d'exemple, citons la résolution, itérative ou directe, des grands systèmes linéaires résultant de la discrétisation de systèmes d'EDP modélisant des phénomènes physiques complexes. La réponse à cette demande passe par la mise au point d'algorithmes adaptés aux plate-formes de calcul parallèles (prise en compte aussi fine que possible de caractéristiques architecturales telles que la hiérarchisation et l'organisation de la mémoire, le réseau d'interconnexion, etc. tout en garantissant la portabilité) et par des développements informatiques s'appuyant sur des couches logicielles capables de gérer ces plates-formes. Ces développements informatiques (incluant des activités de portage sur différentes plate-formes, de validation, de support aux utilisateurs et de documentation) doivent être réalisés par des ingénieurs de recherche et développement, en étroite collaboration avec les chercheurs concernés. La possibilité de recruter de tels personnels est une composante essentielle de la réussite de cette mission. Ce recrutement doit pouvoir se faire sur la base de postes en CDD ou ingénieurs associés (accueil de jeunes diplômés au sein des projets) en charge des principales phases de développement, et de postes statutaires (partagés par plusieurs projets au sein d'une UR) pour garantir la pérennité des développements en question.
- la résolution des grands défis scientifiques posés par notre société nécessitent des actions de recherche et développement multi-disciplinaires associant physiciens, mathématiciens appliqués et informaticiens sur le modèle de celles entreprises aux Etats-Unis dans le cadre du programme SciDAC (Scientific discovery

through Advanced Computing)⁸⁹, s'inscrivant sur une durée de 3 à 5 ans. A cet égard, les ACI (Actions Concertées Incitatives) GRID (Globalisation des Ressources Informatiques et des Données) du Ministère de la Recherche sont de parfaits exemples d'actions multi-disciplinaires dans le domaine des grilles de calcul. Il conviendrait sûrement de s'inspirer des principes à la base de ces actions. Quelques actions de ce type pourraient être initiées chaque année. L'identification de ces actions devrait être de la responsabilité d'un conseil scientifique en charge de la prospective pour le calcul intensif, épaulé par quelques comités de prospectives couvrant un ensemble de thématiques scientifiques identifiées en veillant à ne pas se limiter aux thématiques des comités des centres nationaux (autrement dit, en prenant en compte des disciplines émergentes comme celles discutées à la section 2). Ces actions devraient être dotées de moyens assurant une masse critique suffisante en termes de personnels et permettant une diffusion des résultats dans les principales conférences internationales. En particulier, il serait très souhaitable de flêcher des allocations de recherche en soutien à ces actions ce qui permettrait d'inscrire les travaux dans la durée. Ici, l'INRIA pourrait apporter des contributions significatives sur le plan de l'algorithmique parallèle, numérique ou non, comme sur celui des développements informatiques, à l'image de sa participation active dans plusieurs ACI GRID.

- la recherche française dans le domaine du calcul scientifique intensif gagnerait en visibilité avec l'organisation annuelle de journées nationales dans ce domaine. Il pourrait s'agir de 2 ou 3 journées consécutives, organisées sous l'égide de l'ORAP, articulées autour d'exposés de représentants d'équipes françaises actives dans ce domaine et d'interventions de quelques personnalités étrangères invitées. Ces journées pourraient aussi être l'occasion de présenter en un même lieu l'état d'avancement des travaux menés dans les actions multi-disciplinaires discutées au point précédent. Enfin, chaque instance de ces journées pourrait voir l'édition d'un ouvrage intégrant des articles suffisamment détaillés (de l'ordre de la douzaine de pages pour ne pas se limiter à un survol des activités et résultats), rédigés en anglais et soumis à un comité de relecture.

5 Synthèse

Dans un avenir proche, le calcul intensif ne concernera plus seulement les applications traditionnelles du calcul scientifique (prévisions météorologiques, modélisation de l'atmosphère, de l'océan et du climat, aérodynamique, milieux réactifs, astrophysique, géophysique, etc.) mais verra la montée en puissance de nouvelles disciplines. Nous pensons plus particulièrement aux disciplines qui traitent de problématiques liées à la **modélisation du vivant** (génomique, neurosciences, modélisation du système cardio-vasculaire, etc.) mais aussi à des disciplines plus technologiques (vidéo tout numérique, nanotechnologies, etc.) pour certaines sans relation directe avec le calcul scientifique. L'émergence de ces nouvelles applications s'accompagnera d'une évolution du mode d'utilisation des moyens de calcul intensif : il ne s'agira plus seulement de calculer sur ordinateur mais aussi de **stocker et manipuler de très grands volumes de données**, et d'**interagir sur le déroulement du calcul** via une visualisation quasi-temps réel de son déroulement et la modification de données d'entrée pour des raisons d'optimisation ou d'identification de paramètres. De plus en plus souvent ces grands défis scientifiques seront abordés dans le cadre d'actions impliquant **des équipes et des ressources multi-localisées** (plate-formes de calcul haute-performance, serveurs de visualisation, unités de stockage) . Dans ce contexte, les grilles informatiques répondront au mieux d'une part, aux besoins en capacité de traitement et stockage requises par ces applications et, d'autre part, à la nécessité d'offrir aux chercheurs impliqués des moyens de collaborer activement en dépit d'une multi-localisation géographique. L'INRIA, par son implication forte dans l'initiative GRID'5000, sera à même d'offrir à ses chercheurs un accès à une infrastructure unique. Sa capacité de reconfiguration, notamment en **grille de calcul scientifique intensif**, la variété des équipements, que ce soit pour les aspects calculs, stockage et réseaux, permettront aux chercheurs de réaliser des **expérimentations** dans des conditions excellentes à même de leur permettre d'être au meilleur niveau mondial. Il faudra cependant bien veiller à ce que la pérennité de cette infrastructure soit assurée afin d'offrir aux chercheurs un accès aux technologies les plus avancées. Ce faisant, il

⁸URL :<http://www.scidac.org/>

⁹Voir aussi les rapports de l'initiative américaine SCaLeS (Science Case for Large-scale Simulation)
URL :<http://www.pnl.gov/scales/index.stm>

sera essentiel de préserver le caractère expérimental de cette grille de façon à éviter tout amalgame avec les grilles de production telles que celles qui impliquent les centres nationaux IDRIS et CINES. Clairement, ces deux types de grille ne desservent pas les mêmes besoins et activités scientifiques.

Dans un tel contexte, l'interconnexion entre les ordinateurs joue un rôle central. A cet égard, pour l'INRIA, le grand équipement qui est le plus important, le plus critique pour comprendre l'avenir, est le réseau RENATER. Sa capacité à être un réseau de pointe, permettant outre un fonctionnement de qualité industrielle de l'interconnexion des établissements de recherche et d'enseignement supérieur comme c'est le cas aujourd'hui, mais aussi la capacité à mettre en place des expérimentations réseau véritablement *state of the art* pour des projets comme GRID'5000, nous est cruciale. L'INRIA n'hésite pas en particulier à contribuer au budget de RENATER au delà de ce qu'une analyse purement utilitaire du rapport qualité/prix conduirait à demander, pour être sûr d'avoir accès à cette possibilité.

De nombreux chercheurs de projets de l'INRIA, informaticiens et mathématiciens appliqués, ont collaboré dans un passé récent, ou collaborent actuellement, avec des chercheurs d'équipes extérieures (notamment du CNRS et de l'INSERM) dans le cadre d'actions multi-disciplinaires propres à l'institut (ARC) ou à l'initiative du Ministère de la Recherche (ACI GRID). Les moyens de calcul intensif et les grilles de calcul sont souvent au coeur de ces actions. Les grands défis, actuels et futurs, du calcul scientifique intensif (ou du calcul intensif tout court) nécessiteront des **actions multi-disciplinaires de grande envergure**, s'inspirant par exemple des ACI GRID ou des projets dans le cadre du programme SciDAC, inscrites sur des durées de 3 à 5 ans, associant physiciens, mathématiciens appliqués et informaticiens. Les centres nationaux devront être des participants incontournables des ces actions. Ces actions devront être dotées de moyens suffisants et en particulier, d'**allocations de recherche fléchées**. L'INRIA pourrait sans aucun doute jouer un rôle important dans de telles actions sur la base du savoir faire de ses projets dont les activités touchent au calcul intensif, qui couvre de multiples aspects depuis les couches logicielles les plus basses (systèmes d'exploitation, protocoles réseaux, exécutifs pour les langages parallèles, gestionnaires de ressources, etc.) jusqu'aux applications (du calcul scientifique ou d'autres domaines) en passant par les langages de programmation parallèle, l'algorithmique numérique parallèle, les intergiciels pour l'aide au développement, à l'optimisation et à l'exploitation d'applications de calcul intensif, etc.