

Initiative «Penser le Pétaflop»

Texte pour la journée de restitution des ateliers du 13/11/2008

version du 22/11

Atelier 4-mation

«Les métiers du calcul numérique : formation, recherche et débouchés»

Animateur : Stéphane Cordier et Serge Petiton

Fonctionnement du groupe.

L'atelier 4-mation a fonctionné principalement par des échanges par mail. Il y a eu néanmoins plusieurs rencontres organisées dans le cadre de cet atelier, qui ont bénéficié du soutien d'ORAP

(<http://www.irisa.fr/orap/>). On pourra trouver le programme et les présentations sur les sites web

- 13 juin , journée spécifique, à l'IHP (institut henri poincaré)

<http://www.irisa.fr/orap/Forums/Formation13Juin08/ArchiFormCalcIntens.html>

- 9 octobre sur le site d'EDF Clamart , 23eme forum ORAP

<http://www.irisa.fr/ORAP/Forums/Forum23/programmeF23.html>

L'atelier a fonctionné en 5 sous groupes dont les contributions sont regroupées plus bas.

- 1- Formations LMD, alternance et permanente (Laurent Desbat)
- 2- Definition d'un socle de connaissances adaptatif (Frédéric Louergue)
- 3- Labels sur certaines formations : master, doctorat (Michel Kern)
- 4- Débouchés et (re)valorisation des carrières (Violaine Louvet, Jean-Philippe Nomine)
- 5- Action de communication/motivation/promotion (Marie-Alice Foujols)

Le Haut Conseil pour la Science et la Technologie (HCST) note, dans son [Avis sur la situation de la France en matière de calcul scientifique intensif](#) du 20/12/2007

5. Construire un plan de formation supérieure dans le domaine de la simulation numérique et de son application aux diverses disciplines.

Le retard pris par plusieurs communautés scientifiques françaises dans l'utilisation du calcul intensif et de la simulation numérique doit être rapidement corrigé. Au-delà des actions incitatives qui doivent être poursuivies et développées, **la création d'enseignements à la rencontre des disciplines qui le requièrent et des techniques nécessaires à l'utilisation du calcul scientifique intensif doit être vivement recommandée dans plusieurs établissements d'enseignement supérieur.** Ces enseignements, devraient être labellisés et couvrir les différents cycles d'enseignement supérieur, notamment le doctorat. Dans tous les cas, ils devraient s'intégrer aux enseignements existants, et notamment aux écoles doctorales sans créer de nouvelles spécialisations.

Groupe 1 : Formations LMD, alternance et permanente

Il est essentiel de renforcer la formation à la modélisation et au calcul intensif dans le cadre de la formation initiale LMD pour préparer les futurs diplômés, ingénieurs, chercheurs, dans les entreprises ou dans les organismes de recherche, à l'utilisation des

outils numériques performants (clusters, grilles, HPC, nouvelles architectures multi-coeurs, GPU, outils logiciels associés, etc.). En vue de l'utilisation rapide des machines HPC (celles des jouvences des centres de calcul nationaux, mais aussi celles du projet européen, ainsi que la grille de production en perspective) il faut proposer des formations aux niveaux Doctoral ainsi que des formations permanentes ou en alternance ouvertes aux industriels (et conçues en collaboration avec les industriels).

L'offre de formations est organisée par de nombreux acteurs (universités, centres de calcul nationaux ou mésocentres, chercheurs des organismes à travers les écoles, éventuellement liées à des congrès ou des projets [Grid5000, EGEE, etc.]).

1/ Afin de mieux organiser les formations et de les rendre plus visibles et plus attractives, il serait utile de les recenser, de les lister et de les afficher via un portail unique. Ce recensement pourrait être coordonné par un petit groupe de travail auquel participeraient quelques personnes du réseau des mésocentres (en particulier du projet de groupe Calcul), des centres de calcul nationaux et de GENCI.

2/ Un affichage des formations sur un portail unique serait alors mis en place (avec une bonne visibilité et nécessairement un affichage cohérent des différents niveaux et des différents acteurs). L'analyse du recensement devrait permettre de proposer de nouvelles formations sur les technologies les plus récentes du calcul intensif (multi-coeurs, grilles, GPU, etc.).

3/ Des labels pourraient être associés à ces formations (en lien avec le groupe 3) afin qu'elles soient automatiquement reconnues par les universités comme faisant partie de la formation doctorale (ou permanente).

4/ Une réflexion amont pourrait être menée dans certaines universités avec différentes UFR d'informatique et maths appli. d'une part mais aussi utilisatrices (Physique, chimie, bio, SdU, méca, etc.) afin de mettre au point dans les masters des formations d'initiation à la modélisation numérique, à l'utilisation des machines de calcul intensif, ses méthodes et ses techniques.

Groupe 2 : Définition d'un socle de connaissances adaptatif

Plusieurs rapports dénoncent le retard de la France en calcul de haute performance¹. Les constats convergent et soulignent le manque de moyens de calcul mais aussi le manque de moyens humains. ORAP a ainsi publié un «Plaidoyer pour le calcul haute performance» qui mentionne : La pénurie actuelle de formations universitaires clairement fléchées HPC est très inquiétante. Cette activité doit être clairement reconnue et soutenue.

La définition de socles de connaissances communs disciplinaires a été menée à bien dans différents domaines, comme par exemple le socle de connaissances de la licence de mathématiques², ou des socles liés à la création de PRES. Dans le cas de la définition du socle de la licence de mathématiques, près d'un an de travail a été nécessaire pour aboutir à un texte. L'objectif de cet atelier n'est donc pas de définir le contenu d'un tel socle de connaissances mais plutôt d'en préciser les contours et la méthode pour le définir. Ce socle ne doit pas être un ensemble d'unités d'enseignement mais un ensemble de thèmes qui doivent être traités au cours de la formation.

¹ Académie des technologies, groupe de travail ((Simulation)), Enquête sur les frontières de la simulation numérique, La situation en France et dans le Monde, diagnostics et propositions, Mai 2005 ; Académie des sciences, Groupe d'initiative pour le calcul scientifique, Rapport du 6 Décembre 2006 ; Haut Conseil de la Science et de la Technologie, Avis sur la situation de la France en matière de calcul scientifique intensif, janvier 2008.

² Matapli, numéro 88, p. 35-43

La tâche est ardue puisqu'un socle commun *Calcul intensif* ou *Calcul scientifique haute performance* est nécessairement pluridisciplinaire. Il ne s'agit pas de créer des formations ou d'ajouter des contenus à des formations existantes mais de substituer des contenus *Calcul intensif* à certains enseignements des formations existantes.

Si la formation d'origine est liée à un domaine d'application, ce contenu sera orienté mathématiques et informatique. Si la formation d'origine est mathématique ou informatique, le contenu *Calcul intensif* sera essentiellement orienté vers les disciplines d'applications et les compléments mathématiques ou informatiques.

On peut ainsi imaginer que si des connaissances communes peuvent être dégagées, la priorité qui leur est donnée quant au niveau auquel elles pourraient être acquises, puisse varier. Le socle à un niveau donné peut également varier selon que l'étudiant souhaite ou non poursuivre ses études. Ainsi on peut considérer qu'au niveau licence, le socle de prérequis pour poursuivre en master soit sensiblement différent du socle de connaissances attendus d'un technicien supérieur *Calcul intensif*. Ce socle doit donc être adaptatif.

Groupe 3 : Labels

La préconisation reproduite plus haut du HCST fait écho à des avis similaires, émis par ORAP ou le CSCI sur l'utilité de labels sur les formations. Le contenu de celles-ci sera développé dans le groupe 2, ici nous indiquons quelques points sur la mise en place des labels eux-mêmes.

Il est plus simple de commencer «par le haut», c'est-à-dire par le doctorat. Le cas des masters et licences sera traité ensuite.

Les doctorats

Un label «Calcul Intensif» pour les doctorats pourrait être mis en place assez rapidement. Ce label permettra d'identifier/valoriser les docteurs ayant montré des compétences en calcul haute performance (notion qu'il faudra définir précisément et qui est naturellement pluridisciplinaire, comme précisé dans le rapport du HCST).

Il sera décerné par un comité, dont la composition reste à préciser. Des discussions sont en cours avec différents organismes pour préciser les modalités pratiques de ce projet. Les critères précis d'attribution seront à définir par le comité.

Un point important est qu'il n'existe pas, de tels labels pour les doctorats et c'est l'une des faiblesses qui est souvent mise en avant pour expliquer la difficulté de recruter des docteurs dans le monde industriel (les thèses ne sont pas «lisibles», un tel label serait donc un élément de réponse et cela pourrait avoir un retentissement important, en tout cas, au sein de l'association française des docteurs, l'ANDES, mais aussi peut-être auprès de l'ABG (Association Bernard Gregory). Le calcul intensif jouerait un rôle pionnier.

Ce label mettra en avant le caractère pluridisciplinaire du calcul intensif : à terme, il serait souhaitable (mais ce n'est pas réalisable dans l'immédiat) que les thèses en question soient co-encadrées. Un des encadrants viendrait d'un département de mathématiques ou informatique, l'autre d'un département «applicatif» (à défaut d'un meilleur terme). En ce qui concerne le contenu, il ne suffira pas (sauf

exception) d'avoir développé une nouvelle méthode de résolutions parallèle, ou à l'inverse d'avoir simplement «fait tourner un code sur un super-calculateur». Si le doctorant est mathématicien, ou informaticien, il faudrait qu'il ait pu mettre en œuvre sa méthode sur une application qui intéresse la communauté applicative concernée, s'il vient des «sciences», il lui faudra avoir réalisé une avancée méthodologique.

Dans un premier temps, ce label sera décerné a posteriori, au vu du contenu de la thèse. Mais à terme, il est souhaitable que des prérequis soient mis en place, comme par exemple le fait d'avoir suivi des cours d'école doctorale (qui auront eux-mêmes été labellisés). Ces cours d'écoles doctorales pourraient prendre la forme de formations nationales en calcul intensif, reconnues par les écoles doctorales concernées.

Avoir suivi un master labellisé ne devrait pas être un prérequis (on cherche à élargir la base du public concerné), mais constituerait un avantage.

Une condition indispensable au succès d'un label comme celui-ci est son efficacité, en particulier pour permettre l'insertion professionnelle des docteurs « labellisés ». Cela concerne évidemment l'insertion dans l'industrie, mis également la reconnaissance académique du label (autrement dit, ce label doit avoir une légitimité scientifique suffisante pour offrir un réel avantage au moment des recrutements). Un point connexe est la question de la publicité du label : il doit pouvoir être relayé par les sociétés savantes, et ce dans un grand nombre de disciplines (au-delà des mathématiques et de l'informatique). Une communication vers les industriels du secteur est indispensable, et naturellement vers les associations de doctorants et de docteurs.

Masters, licences

Un label similaire pour les masters et les licences est plus difficile, dans la mesure où il devra nécessairement s'appuyer sur un contenu précis. Ce label devra donc s'appuyer sur le socle de connaissance pour le niveau correspondant.

Une condition nécessaire pour obtenir ce label doit être un réel effort vers la pluridisciplinarité. Si la formation en question est en mathématiques (ou en informatique), elle doit intégrer des enseignements tournés vers des applications (de préférence par des intervenants venant des disciplines concernées), et inversement si la formation est dans un département « applicatif », il est essentiel que des enseignants de mathématiques et d'informatique y interviennent.

Notons que des initiatives similaires existent, comme par exemple la [Designated Emphasis in Computational Science and Engineering](#) de l'université de Berkeley, qui est effectivement un label pour les doctorats. Un programme plus ciblé, au niveau master est celui de [l'Université de Munich](#).

Groupe 4 : Débouchés et (re)valorisation des carrières

Constat

Les filières conduisant à la formation des étudiants dans le domaine du calcul scientifique sont de plus en plus délaissées, rendant le recrutement dans ce domaine difficile. De plus, la quantité de connaissance à acquérir pour former quelqu'un au calcul scientifique augmente de façon constante : les aspects mathématiques, informatiques ... sont de plus en plus complexes et il n'existe pas réellement de filière intégrant l'ensemble de ces éléments. Une formation au calcul scientifique de haut niveau nécessite donc un investissement personnel important. Par ailleurs, les carrières proposées sont d'une part

méconnues, et d'autre part peu valorisées. Enfin, en dehors des entreprises (ou établissements publics) traditionnellement liées au calcul, la simulation numérique est un outil peu connue du domaine privé (notamment des PME).

Propositions

- Communiquer

- En lien avec les recommandations du groupe 5, il est essentiel de mieux faire connaître les métiers du calcul, notamment au niveau master pour inciter les étudiants à s'orienter dans cette voie.

- Les métiers du calcul sont donc peu connus. Pour les valoriser, il faut d'abord les expliquer (sous forme de plaquette, d'un texte ...), en lien avec les propositions du groupe 5. Il faudrait donc envisager avant tout de les décrire, car ils peuvent être très différents les uns des autres, ainsi qu'expliciter les différentes compétences qu'ils requièrent et les besoins auxquels ils répondent. En effet, un aperçu rapide permet d'identifier des métiers assez éloignés les uns des autres :

- développement de codes,
- simulations, analyses,
- assistance, support, aide,
- système massivement parallèle,
- recherche informatique,
- recherche mathématiques appliquées, etc.

Nous recommandons donc, en lien avec les propositions du groupe 5 la mise en place d'un site décrivant les différents métiers du calcul et les différentes filières possibles.

- Dans le même ordre d'idée, il faut également identifier pour mieux communiquer les PME/PMI pour lesquelles l'outil calcul pourrait être une plus-value importante. Une enquête auprès des industriels et certains types de PME pourrait aussi permettre d'évaluer leurs besoins, et leur vision en terme de ressources humaines en rapport avec le HPC. Nous préconisons donc la mise en place d'une enquête sur l'image du HPC auprès des PME/PMI, ainsi que sur leurs besoins potentiels.

- Centraliser les offres de formations, et les offres d'emplois liés au CS et au HPC. Nous recommandons la mise en place d'un site, qui pourrait s'appuyer sur <http://calcul.math.cnrs.fr/>, pour la centralisation des différentes offres d'emplois et des liens associés.

- Valoriser

- Dans le secteur public, l'activité de développement de codes de calcul n'est absolument pas reconnue : il faut que cette activité soit évaluée de la même façon que les publications. L'initiative du projet RELIER (<http://www.projet-plume.org/relier>) va dans ce sens.

- L'importance du rôle des ingénieurs calcul n'est pas (peu) reconnue au sein des laboratoires de recherche et les perspectives de carrières (mobilité, changement de discipline d'applications ...) sont faibles.

Nous recommandons une réflexion devant mener à des propositions précises et concrètes sur la valorisation des publications scientifiques que représentent les codes de calcul, et sur la reconnaissance à sa vraie valeur du travail de développement.

Groupe 5 : Communication

Constat :

Le HPC et les métiers associés sont méconnus du grand public, or nous avons un besoin en personnel motivé et compétent. Ce paragraphe décrit quelques propositions afin de

mieux faire connaître cette discipline, et les métiers associés, avec le but avoué d'encourager les vocations auprès des jeunes.

Propositions :

- Site web, orienté jeunesse (plusieurs niveaux : élémentaire, collège, lycée, supérieur)
 - o Description du calcul haute performance
Mise en ligne, mutualisation d'exposés de niveaux différents
 - o Exemples de réalisations
 - o Liens avec les disciplines scolaires : mathématiques, informatiques, physique, chimie, géographie
www.nasascience.nasa.gov
<http://interstices.info/>
<http://www.animath.fr/spip.php?rubrique6>
 - o Soutien aux projets scolaires type TIPE
 - o Liens vers les formations (et les labels), le socle des connaissances, les débouchés (groupes 1, 2, 3 et 4)
 - o Fiches métiers, chiffrage des débouchés (exemple des brochures « zoom sur les métiers de... » fait en partenariat avec l'ONISEP)
 - o Les partenariats européens et internationaux
 - o Les sites voisins :
www.imaginetonfutur.com/ recense tous les métiers
asso.objectif-sciences.com/ organise des animations scientifiques : vacances scientifiques, conférences, ...
- Site web orienté presse :
 - o Exemples de réalisations
 - o Films et animations
 - o Fiches descriptives
- Participation aux conférences de vulgarisation scientifique et communications en direction du grand public (exemple des suppléments du magazine « la recherche » fait en collaboration avec le CEA de janvier 2006 et mai 2007)
- Site web orienté communauté HPC (existent déjà : ORAP : www.irisa.fr/orap , <http://calcul.math.cnrs.fr/>)

Le centre national Jacques-Louis Lions ou maison de la simulation:

Il est nécessaire de diffuser et renforcer le message en mutualisant les actions déjà menées dans ce domaine et en leur donnant explicitement un but commun : faire connaître le calcul haute performance et encourager les vocations. Le CEA, le CNRS, ORAP et les acteurs majeurs du calcul scientifique de haute performance en France (Universités, EDF, Météo France, Total, Académie des sciences, Académie des technologies, ...) seraient naturellement encouragés à rejoindre une initiative fédérant ces actions.

Quelques acronymes utilisés

CPU : Conférence des présidents d'universités

CS : Calcul Scientifique

CSCI : Comité Stratégique pour le Calcul Intensif

HCST : Haut comité pour la Science et la Technologie

GENCI : Grand équipement national pour le calcul intensif

HPC : calcul haute performance

LMD : Licence Masters, Doctorats,

Recommandations

1_ Socle de connaissance "CI" (master, licence)

Que ce soit niveau licence (prérequis pour un master) ou en master (prérequis pour un doctorat), il y a un consensus général pour dire que ce serait très utile mais que c'est une tâche ardue (en raison notamment du caractère multidisciplinaire et évolutif du CI). Il n'y a pas de candidat déclaré pour proposer un tel socle qui nécessiterait un accord de l'ensemble des acteurs concernés.

2_ Labels (doctorat, master)

Niveau D : un projet est en cours, porté par GENCI et soutenu par la CPU, il a été annoncé le 9/10. Il s'agit de labeliser les docteurs a posteriori (après leur soutenance).

Niveau M : il s'agirait cette fois de labeliser les formations. Tant qu'un socle n'est pas défini, ce n'est pas facile d'en définir les critères et on doit également connaître l'existant (point suivant). 2 conditions nécessaires pourrait être utilisées : d'une part que la formation s'appuie sur des modules ou projets utilisant une machine parallèle, auquel les étudiants auraient accès, typiquement un méso-centre et d'autre part, prendre en compte la pluridisciplinarité. Dans la trilogie Math, info et domaine d'application, on pourrait par ex. exiger qu'au moins 2 de ces trois items soient représentés significativement dans les modules suivis obligatoirement par les étudiants.

3_ Recenser toutes les formations

Il serait utile de connaître l'état des lieux, de l'ensemble des formations (à partir du master , mais également les cours d'écoles doctorales, les formations organisées par les méso-centres ou les centres nationaux, les écoles spécialisées). Ce recensement a été initié par le groupe calcul, il faudrait qu'il soit le plus exhaustif possible, ce qui demande un soutien de l'ensemble des acteurs

4_ Recenser les besoins industriels (notamment PME)

Une enquête en ce sens serait utile et permettrait de mieux identifier les besoins à venir et donc les débouchés possibles (point 5). La mobilisation des PME sur ces questions est difficile et cela doit s'inscrire dans le moyen terme. Certains éditeurs de logiciels ont commencé à anticiper une évolution sur ces questions. Il n'y a pas de candidat déclaré pour ce recensement.

5_ Faire connaître les métiers du numérique

Ce serait très utile et est lié aux points 3 (connaître les formations et leurs débouchés), 4 (connaître les besoins à venir des industriels) , 6 (connaître l'offre d'emploi actuelle) et 7 (faire connaître les opportunités d'emploi dans ce secteur). Une action de type brochure ONISEP "Zoom sur les métiers du CI" pourrait être un objectif , susceptible d'associer la plupart des acteurs du CI. Il n'y a pas de candidat déclaré pour ce coordonner cette action.

6_ Centraliser les offres d'emploi

Une action dans cette direction est en cours par le groupe calcul (site web et liste de

diffusion à laquelle sont déjà inscrit plus de 600 personnes). Elle pourrait être amplifiée et relayée par l'ensemble des acteurs.

7_ Communiquer vers les jeunes, les médias

Compte tenu de la désaffection des filières scientifiques, ce point est essentiel pour le moyen terme. De telles actions nécessitent des moyens importants, notamment humains, et pérennes. Il n'y a pas de candidat déclaré pour ce coordonner cette action.

8_(Re)valoriser les carrières, primes

En ce qui concerne, les doctorants, notons l'initiative, annoncée lors de l'inauguration de la machine du CINES, des prix « master de ParisVI (ex-prix 421) » par Yvon Maday.

Les universités auront bientôt des moyens accrus pour encourager des projets multidisciplinaires, constituer des équipes importantes, susciter des vocations... via les réformes du système de prime et de promotions. Il faudrait qu'elles soient sensibilisées à l'importance de ces questions et également trouver des moyens pour faciliter la mobilité entre les différents métiers (ingénieur, chercheur et enseignants chercheurs).

Les organismes pourraient soutenir de tels projets pluridisciplinaires par exemple par le système de délégation d'enseignant-chercheur sur plusieurs années.

9_ Proposer une action type « scidac » européen

Une initiative en ce sens est envisagée au sein d'ORAP.

Annexe : Liste des inscrits à la liste de discussion de l'atelier

stephane.cordier@math.cnrs.fr, Serge.Petiton@lifl.fr, Ronan.Keryell@hpc-project.com, vialle@metz.supelec.fr, Chantal.Letonqueze@irisa.fr, Francoise.Berthoud@grenoble.cnrs.fr, philippe.depondt@insp.jussieu.fr, hinsen@cnrs-orleans.fr, claude.cohen@bull.net, alain.kavenoky@cea.fr, annie.spielfiedel@obspm.fr, nataf@ann.jussieu.fr, leclerc@lmt.ens-cachan.fr, Fabienne.Jezequel@lip6.fr, farge@lmd.ens.fr, Jean-Marc.LAYMAJOUX@edf.fr, Laurent.DADA@cea.fr, Jean-Philippe.NOMINE@cea.fr, Yvon.maday@ann.jussieu.fr, alain.samblat@dassault-aviation.com, laurent.desbat@imag.fr, Jean-Luc.Lamotte@lip6.fr, cedric.cocquebert@cea.fr, quang.dinh@dassault-aviation.fr, Christian.Lucius@dassault-aviation.com, gupta@lri.fr, michel.kern@recherche.gouv.fr, stratis.manoussis@cnrs-dir.fr, gbarat@sgi.com, Nahid.Emad@prism.uvsq.fr, philippe.thierry@intel.com, jpo@proba.jussieu.fr, thual@imft.fr, marc.till@airliquide.com, guillaume.poulain@cggveritas.com, patrick.wohlschlegel@fr.ibm.com, marc.massot@em2c.ecp.fr, cheikh.diop@cea.fr, sibilla@ipgp.jussieu.fr, Jean-Claude.Andre@cerfacs.fr, Jean-Loic.Delhaye@inria.fr, Stephane.Lanteri@sophia.inria.fr, jy.berthou@edf.fr, mark.asch@u-picardie.fr, david@icps.u-strasbg.fr, Marie-Alice.Foujols@ipsl.jussieu.fr, louvet@math.univ-lyon1.fr, adlane.sayed@univ-artois.fr, senechaldorothee@gmail.com, frederic.loulergue@univ-orleans.fr