

ASTROSIM

Le colloque AstroSim (<https://astrosimconf.sciencesconf.org>) qui a eu lieu à Lyon sur 4 journées, du 8 au 11 octobre 2018 a rassemblé 64 personnes : chercheurs, ingénieurs de recherche, post-doctorants et étudiants venant de toutes les thématiques astrophysiques, représentants institutionnels (directeur du CINES, président du Comité Thématique Astrophysique et Géophysique, directeur de la maison de la simulation), ou acteurs privés (VP Atos-BULL en charge du HPC).

Au delà des résultats scientifiques, les orateurs ont tous donné quelques éléments de réflexion prospective, sur les enjeux stratégiques associés aux activités de simulation numérique ou sur les difficultés persistantes qu'ils rencontrent dans l'exercice de leur métier. Trois sessions plénières de discussion ont complété ces remarques. Nous donnons ci-dessous une transcription de ces trois discussions, et proposons une synthèse.

Résumé

Les codes de simulations numériques sont devenus de grands instruments à part entière, qui sont déployés sur la TGIR GENCI, sur les mésocentres, et sur les centres de calculs internationaux. Le développement et l'exploitation de ces instruments numériques demande aujourd'hui un niveau de professionnalisme et une diversité d'expertises tout à fait comparables aux besoins des grands instruments d'observation. Pour répondre aux enjeux auxquels fait face la communauté des simulations numériques en astrophysique, nous faisons trois recommandations :

La **création d'une structure AstroSim** (par exemple une *Action Spécifique*) avec une mission triple: (1) animer et former notre communauté, (2) structurer notre communauté et éventuellement coordonner des actions de développement majeures, (3) représenter nos intérêts auprès de nos tutelles et partenaires.

Le **recrutement et le redéploiement d'ingénieurs de recherche** pour le développement des codes de simulation numérique et des outils d'analyse associés, à la fois au plus près des équipes de recherche et en partenariat avec des centres d'expertise comme la Maison de la Simulation.

La **labellisation par l'INSU de certains codes communautaires**, dans un ANO 2.3 : *instrumentation numérique pour les centres de calcul.*

NB: Ce texte a été rédigé par J. Blaizot (coord.), B. Commerçon, B. Dintrans, P. Hennebelle, G. Lesur, E. Millour, et P. Tremblin. Il a été approuvé par l'ensemble des participants du colloque AstroSim 2018.

Discussion #1 : Les moyens matériels

Centres nationaux

La communauté astro est très présente au niveau des centres de calcul nationaux : nous sommes parmi ceux qui obtenons le plus de temps et produisons le plus de données. Notre communauté est globalement satisfaite mais nous notons les points suivants.

- Il semble que les utilisateurs ne soient pas toujours conscients de l'aide que ces centres peuvent leur apporter pour déployer ou optimiser (dans une certaine mesure) leurs applications. Nous encourageons les chercheurs à utiliser davantage ces ressources.
- Nous notons qu'il y a un facteur 2 à 3 entre la puissance de calcul offerte aux chercheurs allemands et français. La dernière campagne d'attribution de temps de calcul GENCI s'est faite avec un taux de pression allant jusqu'à dépasser 400% sur les machines généralistes.

Mésocentres

Le parc de mésocentres est assez hétérogène en puissance de calcul disponible, de support aux utilisateurs, et de lien avec les communautés astro locales. Il y a une certaine inquiétude de notre communauté par rapport à la pérennité des moyens de calcul offerts par les mésocentres. Le financement de nombreux mésocentres par l'IDEX Equip@meso en ~2011 a permis un effet de levier temporaire, mais nombre de structures n'ont aujourd'hui plus de financement pérenne à l'heure où le renouvellement des calculateurs à l'échelle régionale s'annonce. Dans ce contexte, on peut se demander s'il est plus pertinent d'investir dans des moyens locaux (UMR/OSU) que dans les mésocentres (par exemple quand on obtient quelques centaines de k€ via une ERC ou autre initiative d'excellence). Ce raisonnement néglige généralement l'infrastructure et l'expertise que ces mésocentres ont développées, qui semblent pérennes, et qui permettent une mutualisation efficace des coûts entre plusieurs partenaires. Ces coûts sont trop élevés pour être portés par les UMRs en général et le sentiment général est que les mésocentres sont une solution qu'il faut continuer de développer. En effet, une évolution possible des Tier-2, si leurs moyens de calcul ne sont pas renouvelés, est qu'ils proposent des services d'hébergement mutualisés (par exemple des serveurs achetés avec des ANRs, etc., ou en facturant les heures pour ceux qui ont obtenu des crédits pour cela) et des services de soutien au développement. Certains centres (comme GRICAD) ont déjà mis en place avec succès un tel système de mutualisation. Il nous semble que l'INSU doit soutenir les mésocentres, au moins comme solution partielle à la crise du Tier-3, en encourageant les OSUs et UMRs à s'en rapprocher, et à y mettre des moyens humains et matériels. Il faut noter que cet enjeu dépasse bien sûr notre communauté simulations numériques.

Notons qu'un point de blocage récurrent est la difficulté d'inclure du matériel financé par l'ERC dans les mésocentres, ce qui les prive d'une source de renouvellement importante (et met en difficulté les unités de recherche dont l'infrastructure n'est pas forcément dimensionnée pour héberger de tels serveurs). Un exemple qui semble marcher consiste à inclure dans la demande ERC un paragraphe expliquant que le groupe ERC a priorité sur la machine pour un volume d'heure correspondant au montant de la contribution de l'ERC. Par exemple (ERC de J-L Barrat pour une contribution à CIMENT) :

Use of computing facilities:

The host institution offers the possibility to host the computing facilities within the framework of the "ciment" (Calcul Intensif, Modélisation, Expérimentation Numérique et Technologique, <http://ciment.ujf-grenoble.fr>) program of the University Joseph Fourier, which will provide for no charge the necessary infrastructure in terms of power and cooling, and basic technical support. The team will also benefit from the offers negotiated by CIMENT with High performance computing vendors. In counterpart, the participants to CIMENT will be allowed to use the computing infrastructure of the team in the so called "best effort" mode, which only submits low priority tasks that are suppressed as soon as the facility is required for the use by a member of the team. The computing times used by the team members and in best effort mode by participants of the CIMENT project will be accounted for separately. However, as the team members have fully priority access to the computing facility funded by the project, the equipment will be fully charged to the project provided a reasonable part of the time is used by team members.

Evolutions

Il y a de fortes inquiétudes par rapport aux architectures futures que proposeront les centres: seront-elles adaptées aux codes de la communauté ? Et sinon comment se préparer efficacement à ces nouvelles architectures étant donné les faibles moyens humains. Le président du CT4 et le directeur du CINES encouragent notre communauté à s'exprimer plus fortement sur ses besoins pour pouvoir peser dans les choix matériels qui seront faits et les conditions qui les accompagneront. Il n'y a pas réellement de canal officiel pour ce type de communication, et la constitution d'une entité qui représente notre communauté permettrait d'avoir une voix plus forte. Il est important de noter qu'il ne s'agit plus d'évolutions à long terme : ces évolutions sont déjà en place et les besoins d'adaptation des codes sont urgents (par exemple sur Irene Skylake, censée être 2.5-3 fois plus rapide que Curie, on n'arrive qu'à obtenir un facteur ~1.5 au mieux).

On voit se dessiner une grosse problématique en ce qui concerne les données. Produites en masse sur les centres nationaux, il n'est parfois pas envisageable de les rapatrier ailleurs ; à moins de développer des mésocentres plus orientés "données" que "calcul" ? Mais ces derniers demanderaient toutefois de résoudre le problème du transfert de ces données pour archivage et post-traitement. Si ces données sont amenées à rester sur les centres nationaux, alors il faut d'ores et déjà s'y préparer. Le CINES s'engage dans cette voie et proposera en plus du HPC des services / plateforme pour l'analyse intensive de données (HPDA).

Divers

Il n'est pas toujours facile pour un utilisateur (ou groupe) de savoir s'il doit s'adresser à un centre national ou à un mésocentre : certaines allocations de temps obtenues sur des centres nationaux auraient plus leur place sur des mésocentres, et inversement. Une coordination entre les centres nationaux et les mésocentres serait souhaitable pour harmoniser les attributions. Une initiative pilote menée dans ce sens par le CINES en partenariat avec quelques mésocentres (CALMIP, GRICAD, ...) est en train de voir le jour. Faire mieux connaître aux chercheurs les moyens accessibles et leurs contours est aussi l'un des enjeux des écoles et colloques AstroSim.

Discussion #2 : Les moyens humains

Les codes de simulations numériques sont devenus des instruments à part entière et demandent aujourd'hui un niveau de professionnalisme que notre communauté peine à mettre en place. Plusieurs points ressortent des discussions :

- Le besoin du soutien d'Ingénieurs de Recherche pour le développement et le maintien de codes est exprimé de manière quasi unanime.
- Certaines unités ont su mettre en place en interne un support efficace au développement HPC, par exemple via des appels d'offre réguliers. Tous les exemples qui ont été évoqués dans ce sens sont des succès. La proximité géographique et thématique des IRs apparaît comme un élément clef de ce succès. Il semble important de ce point de vue qu'ils soient placés dans les UMRs.
- Les partenariats avec les mésocentres ou centres d'expertise sont relativement nombreux et là aussi souvent positifs. Encore une fois, on note le manque d'IRs dans les unités qui permettraient de donner plus de résonance à ces échanges.
- Des nouvelles architectures apparaissent déjà dont nous avons du mal à tirer profit. Nous avons besoin de compétences pour cela.

Une partie importante de la discussion a porté sur la meilleure manière pour notre communauté d'exploiter les moyens qui nous sont offerts via les centres d'expertises comme la Maison de la Simulation. Des thèses en co-tutelle sont (étaient ?) proposées mais paraissent parfois difficiles à mettre en oeuvre (au moins pour les unités hors région parisienne) et sont finalement peu utilisées. Plusieurs freins peuvent être identifiés dans notre relativement faible utilisation de la MdS:

- L'absence d'appel d'offre : la manière dont la MdS choisit ou non de s'engager sur des projets de développement est opaque. Ajoutant à cela que la MdS est déjà engagée à plus de 100% sur de nombreux projets, il est difficile pour un chercheur isolé d'obtenir des ressources significatives.
- Le temps long des développements : Porter un code pour l'exascale veut souvent dire tout ré-écrire en repensant la structure profonde du code. Ce type de développement est long (plusieurs années) et coûteux. La MdS n'a ni la vocation ni les ressources pour porter seule de tels projets et il est difficile pour les chercheurs de se lancer dans cette direction sans soutien IR relativement pérenne.
- L'absence d'une stratégie lisible de notre communauté, et d'une volonté politique forte pour la mettre en oeuvre, est un handicap majeur pour développer des partenariats sérieux avec les centres d'expertise.

Concernant l'octroi de postes d'ingénieurs, il est déploré que "la situation d'urgence instrumentale" conduise bien souvent, au niveau des UMRs, à reléguer les numériciens en seconde priorité et à repousser les embauches parfois de plusieurs années. Or cette situation, même si moins "visible" que pour un instrument, n'en est pas moins dommageable. En particulier, elle freine de manière très préjudiciable les développements numériques de grandes ampleurs.

Discussion #3 : Les codes communautaires

Partant du panorama des divers codes utilisés par la communauté astrophysique nationale, on ne peut que constater l'extrême diversité des outils numériques mis en oeuvre avec des degrés divers de maturité. Il apparaît donc nécessaire de mieux cerner les outils susceptibles de devenir "communautaires" à travers certains critères :

- Un nombre critique de développeurs de l'outil (parallélisation, portage sur les nouvelles architectures, structure logicielle, algorithmes, etc)
- Ou un nombre critique d'utilisateurs de l'outil (ajout de nouvelle physique, production de données simulées)
- Ou un nombre critique d'utilisateurs des données produites par l'outil (post-traitement, comparaison avec les observations)

Si l'un de ces critères est rempli, il apparaît nécessaire aujourd'hui de mettre en place des actions pour le support à la communauté sur les points suivants correspondants :

- Développement du code : support pour le portage sur les architectures Exascale. Les codes doivent e.g. exploiter >100,000 coeurs avec ~1Go de RAM/coeur efficacement. L'utilisation de nouveaux algorithmes plus adaptés aux nouvelles architectures massivement parallèles est aussi une piste à privilégier.
- Utilisation du code : support pour la maintenance et la pérennité des codes. Il est nécessaire de mettre en place une stratégie pour assurer la pérennité de ces codes, notamment pour l'ajout de nouvelles physiques par des interfaces "user friendly". Il est aussi nécessaire autant que possible de faire grossir la communauté d'utilisateurs, en assurant des formations longue durée (1-2 mois) des utilisateurs (étudiants, postdocs, chercheurs) sur des codes communautaires bien ciblés. Il est aussi nécessaire d'éduquer la communauté sur les licences libres (Cecill, GNU GPL) et expliquer les risques des [licences informelles](#).
- Utilisation des données : support au stockage, à l'accessibilité et à la ré-utilisation des données produites. Il est nécessaire de fournir des bases de données libres d'accès avec des outils permettant une ré-utilisation rapide des données, notamment pour faciliter la comparaison avec les observations.

En parallèle de ce travail sur les codes "communautaires", il est aussi souhaitable de permettre/soutenir l'émergence de nouveaux codes et éviter qu'ils soient moins soutenus que des codes déjà établis. A cette fin, une action de soutien pour l'émergence de nouvelles approches serait opportune et il apparaît nécessaire qu'un label "communautaire" soit ré-évaluable et ne soit attribué que pour une période limitée dans le temps (tout en étant reproductible).

Lors de ces discussions, une liste (non-exhaustive) de codes a été mentionnée: [RAMSES](#), [Meudon_PDR](#), [Einstein Toolkit](#), [ESTER](#), [MagIC](#), [Lorene](#), [Kadath](#), [PLUTO](#), [DYNAMICO](#), [LMDz](#), FARGOCA, AMRVAC.

Conclusions

De ces trois discussions, et de l'ensemble des interventions, nous tirons les conclusions suivantes.

La communauté simulations numériques astrophysiques demande la création d'une structure AstroSim (par exemple une action spécifique). Cette entité devra répondre à plusieurs besoins avec les missions suivantes:

1. Animation et formation de la communauté.
 - a. Organisation de rencontres régulières (colloques AstroSim) pour soutenir des échanges fréquents, tant sur les méthodes numériques que sur les questions scientifiques ou les pratiques, dans notre communauté.
 - b. Organisation d'écoles régulières (écoles AstroSim) dans lesquelles sont proposées des formations aux techniques numériques astrophysiques et en général (méthodes numériques, parallélisme, HPC et HPDA,...), et des formations pratiques sur les codes de notre communauté.
 - c. Faire connaître les moyens (internationaux, nationaux et régionaux) auxquels ont accès les chercheurs. Les écoles et colloques AstroSim sont aussi l'occasion de faire rencontrer à notre communauté les différents acteurs publics ou privés du HPC en France (et réciproquement). Par exemple, les écoles AstroSim sont hébergées dans des centres de calcul (Tier-2 ou Tier-1) : le Centre Blaise Pascal à Lyon en 2017, le CINES en 2019.

2. Structuration de la communauté.
 - a. Identifier les codes communautaires ou qui ont vocation à le devenir et accompagner la structuration des communautés d'utilisateurs / développeurs pour obtenir une labellisation par l'INSU.
 - b. Dans certaines disciplines, aider l'émergence d'un code à vocation communautaire parmi une multitude de codes similaires.
 - c. Soutenir l'émergence de nouveaux codes. Il est important que le soutien aux codes communautaires ne fasse pas disparaître le foisonnement d'initiatives originales ni ne dévalue les codes qui n'ont pas vocation à être utilisés par une large base d'utilisateurs.
 - d. Mener une action de veille technologique et méthodologique et encourager des initiatives de collaborations interdisciplinaires ou intercommunautaires qui sont souvent longues à donner des fruits.
 - e. Mener une réflexion stratégique globale sur la prochaine génération de codes (fonctionnalités, développements techniques spécifiques, adéquations aux machines exascales à venir à l'horizon de 5-10 ans) *et la mettre en oeuvre*. Un certain nombre de pistes ont été identifiées et doivent être creusées, comme le rassemblement autour de librairies communes de maillage adaptatif massivement parallèle, et l'écriture d'un *livre blanc* décrivant en détail les objectifs d'un (ou plusieurs) nouveau(x) code(s) communautaire(s)

est une action à mener rapidement pour éventuellement lancer des développements communs coordonnés.

3. Jouer le rôle d'interlocuteur auprès des différents partenaires et tutelles.

- a. Informer GENCI (et plus généralement les centres de calculs de Tier 0, 1, 2) des besoins de notre communauté en termes d'architectures (mémoire, parallélisme, etc.) de sorte à ce que ces besoins soient pris en compte dans les choix matériels à venir. Cela est particulièrement important dans un contexte où de nouveaux acteurs émergent avec des besoins radicalement différents (e.g. intelligence artificielle).
- b. Proposer à la CSAA des avis motivés pour des questions de prospective, de stratégie, de labellisation, de classement de priorités relatives aux simulations numériques.
- c. Faire remonter à l'INSU et MiCaDo les spécificités de nos besoins, et faire redescendre à notre communauté les évolutions institutionnelles. La réflexion collective menée en 2.e pourrait mener vers des propositions de postes IR portés par l'INSU et affectés par MiCaDo par exemple à la Maison de la Simulation.
- d. Définir avec la Maison de la Simulation le contour d'un partenariat qui pourrait aller du conseil à un soutien réel au développement selon les moyens que nous y mettons.

On note que certaines actions sont limitées dans le temps par nature (e.g. 2.e, 3.d) mais d'autres ont vocation à s'inscrire dans la durée.

Les codes de simulation sont des instruments numériques dont le développement, le maintien, les évolutions, et l'exploitation requièrent le soutien d'ingénieurs de recherche. Les expériences qui marchent nous montrent qu'il faut des IRs dans les UMR, proches des équipes de recherche, et qui partagent en partie leur culture thématique. Ces IRs sont nécessaires pour faire le lien avec les expertises proposées dans les centres de compétence (e.g. la Maison de la simulation, Maimosine, ...) et les chercheurs, et ainsi amener de nouvelles méthodes dans notre communauté. L'INSU doit aussi se donner les moyens de renforcer ces centres interdisciplinaires pour qu'ils puissent dégager du temps pour répondre à nos besoins. Cela peut se faire via MiCaDo ou bien par des accords locaux (par exemple à GRICAD, des personnels techniques (IR, IE) de différentes unités font une partie de leur service au mésocentre), par exemple pour mener des développements d'intérêt général identifiés par la structure AstroSim (point 2.e plus haut).

Certains codes communautaires devraient être soutenus par une labellisation de l'INSU. Par exemple, un code tel que RAMSES semble déjà mûr dans la mesure où il est utilisé par une vaste communauté d'utilisateurs issus de thématiques très diverses. Un code de cette nature pourrait porter un tel projet. Cette labellisation est typiquement de l'ordre

d'un ANO2.3 : *instrumentation numérique pour les centres de calcul*. Elle vise à répondre aux besoins suivants:

- Structurer la communauté utilisatrice et développeuse du code. Notamment, faire apparaître une gouvernance qui puisse se prononcer sur les évolutions du code (intégration et homogénéisation de développements menés par différentes équipes, choix de conventions, directions stratégiques à soutenir, ...). Cet aspect est particulièrement important dans le cadre de codes internationaux.
- S'assurer de la maintenance du code et de sa documentation.
- Rendre le code plus accessible, en assurant des formations pour utilisateurs ou développeurs, à la fois avec des écoles dédiées et par un guichet de support.
- Organiser des ateliers annuels pour animer la communauté utilisatrice.
- Assurer la pérennité du soutien ingénieur.