

Forum Teratec : omniprésence du calcul numérique

Du 1^{er} au 2 juillet 2014, l'École Polytechnique de Palaiseau accueillera le Forum Teratec dédié à la simulation numérique et au HPC (calcul haute performance). Deux jours de conférences et d'ateliers thématiques autour d'un domaine incontournable pour la compétitivité des entreprises françaises de nombreux secteurs d'activité.

Il semble que rien ne puisse échapper au calcul numérique et aux modèles mathématiques inventés pour modéliser virtuellement le monde qui nous entoure. Météorologie, finance, biologie, astrophysique, industrie mécanique, géologie, cinéma... autant de domaines gagnés par le calcul numérique. Quelle que soit la complexité d'un système, il devient en effet possible de simuler virtuellement son fonctionnement dans son environnement réel. Les objectifs sont multiples : mieux comprendre les lois qui le régissent, améliorer ses performances, prédire son évolution, générer des images d'animation réalistes, calculer des niveaux de risques, inventer de nouvelles solutions, etc.

Chercheurs et ingénieurs font donc davantage de calculs, sur des modèles toujours plus imposants et bien entendu ne supportent plus d'attendre des heures les résultats de leurs simulations. Cette inflation de moyens a un nom : HPC pour High Performance Computing, et une rencontre annuelle, le Forum Teratec, organisée par l'association éponyme. En moins de dix ans, ce dernier



Le calcul numérique est désormais présent dans tous les domaines d'activité humaine.

s'est imposé comme l'événement incontournable du petit monde du calcul haute performance. Les organisateurs attendent cette année un millier de professionnels à l'École Polytechnique de Palaiseau. Au menu de ces deux journées, les habituelles sessions plénières,

les ateliers techniques et une exposition des principaux fournisseurs du secteur.

Une édition 2014 qui sera sans doute plus attendue que les précédentes, étant donné l'élan insufflé cette année par le ministère du redressement

Bull recrute 500 ingénieurs en France

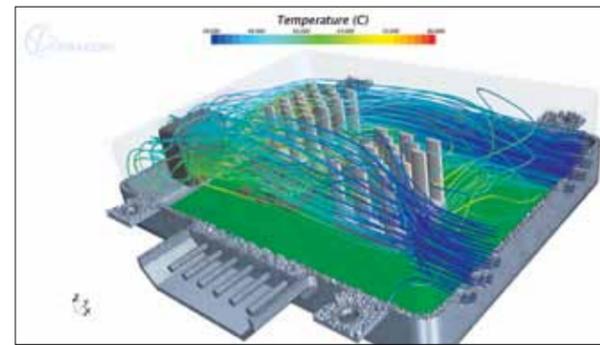
Le constructeur français souhaite recruter 1000 ingénieurs, dont 500 en France dans ses domaines de prédilection : calcul haute performance, Big data et Cloud Computing. Bull cherche plus précisément des ingénieurs Cloud, HPC, Sécurité, Développement, Business Intelligence, Systèmes et R&D.

productif à travers ses « 34 plans pour une nouvelle France industrielle ». Car, parmi ceux-ci figurent en bonnes places le Big Data, le Cloud Computing, et les supercalculateurs. Trois domaines imbriqués, identifiés comme stratégiques pour la compétitivité française. Signalons que l'un d'entre eux, le plan « Supercalculateurs » est dirigé par Gérard Roucairol, le président de Teratec (voir page 21).

HPC : marché et enjeux

Si l'on considère l'ensemble du marché mondial, le HPC (systèmes, stockage, logiciels, applications et services) représentait selon IDC 14 milliards d'euros en 2010. En 2010 et 2011, le secteur a enregistré des croissances respectives de 10 et 8,4 %. La croissance moyenne observée jusqu'à maintenant est de 7,5 % par an ; elle devrait afficher au moins le même score dans les trois ans à venir.

Bien logiquement, ce sont les grands laboratoires de R&D et les grandes entreprises comme EDF, Areva, Airbus, Total ou Intel et Google qui ont les moyens de se payer des supercalculateurs de plusieurs millions d'euros sans compter les frais de maintenance. Leurs besoins en la matière ne cessent de croître. L'un des enjeux du HPC est donc de fournir une puissance de calcul toujours plus importante. On évoque la construction de machines exaflopiques d'ici 2020, c'est-à-dire capables d'effectuer un milliard de milliards d'opérations à la



Dans le domaine industriel, la mécanique des fluides est particulièrement gourmande en ressources calcul et bénéficie à plein de l'utilisation de machines HPC massivement parallèles.

seconde. Comme repère, le Très Grand Centre de Calcul (TGCC) du CEA à Bruyères-le-Châtel offre une puissance de calcul dépassant le pétaflop, soit 1 million de milliards d'opérations par seconde.

Mais, le second enjeu est la diffusion de ces ressources technologiques vers les PME, industrielles notamment. Et là, l'affaire se complique. Pour preuve, les programmes comme « Initiatives HPC-PME » ou HPC Luminov qui proposent le Cloud comme moyen d'accès aux outils de calcul ont du mal à prendre. Car, si les PME peuvent avoir des besoins semblables aux grands groupes, leurs ressources sont limitées sur le plan financier, mais aussi sur le plan de la compétence. « Elles doivent donc être accompagnées de services adaptés pour être en mesure d'intégrer ces outils dans leurs processus de conception et d'industrialisation » souligne Brigitte Duème responsable du transfert technologique Aéronautique, Spatial/Défense, Sécurité, Robotique et Calcul Intensif à l'Inria (à l'origine d'Initiatives HPC-PME). Pour l'instant, le bilan de

l'association reste modeste. Six ou sept entreprises ont mené au bout leur projet, une vingtaine sont encore en phase d'études. « Ce type d'approche est très innovante et se heurte donc à beaucoup de difficultés

Travailler à l'échelle européenne

Il semble cependant que pour remporter ces enjeux, l'échelle nationale soit déjà largement dépassée. Les entreprises françaises du secteur ne peuvent se battre seules face à des concurrents nettement plus armés



Les besoins en terme de puissance ne cessent de croître. L'objectif est d'atteindre d'ici 2020 des machines exaflopiques (un milliard de milliards d'opérations à la seconde).

financières, organisationnelles, connaissance du sujet des entreprises, etc. Mais, la preuve est faite qu'elle fonctionne et peut permettre à des PME de faire leurs premiers pas dans le calcul numérique de haut niveau pour gagner en perfor-

comme les USA ou la Chine qui lancent tous deux des plans gouvernementaux massifs pour développer leurs champions en matière de HPC. Pour la Commission Européenne, l'avenir du secteur HPC en France passe obligatoirement par une



Forum Teratec : demandez le programme

- **Mardi 1^{er} juillet** : Les sessions plénières seront centrées sur les défis technologiques de la simulation numérique haute performance et sur la diversité des usages du calcul intensif. On pourra assister à des présentations faites par les responsables d'Airbus, de la plateforme européenne ETP4HPC, du projet Square Kilometre Array (SKA), de UberCloud HPC Experiment, de Distene, etc. Arnaud Montebourg est également attendu pour la présentation du Plan Supercalculateurs qui fait partie des 34 plans de reconquête de la « Nouvelle France Industrielle ».
- **Mercredi 2 juillet** : Des ateliers techniques orientés « technologies » ou « usage » seront animés par les principaux acteurs du marché qui feront le point sur les technologies émergentes et sur de nouveaux secteurs d'application du HPC.

Côté technologie seront abordés quatre sujets :

- Évolutions des applications et des architectures HPC & Big Data – Influences, confluences.
- Conception numérique optimale des systèmes complexes : état de l'art et verrous technologiques.
- Calcul scientifique & Open Source : pratiques industrielles des logiciels libres.
- HPC & Big Data – Logiciels et outils.

Pour les ateliers usages :

- Simulation des matériaux : nouveaux matériaux, usage et tenue en fonction.
- Technologies numériques pour le végétal.
- HPC et Santé : De la recherche thérapeutique à la médecine personnalisée.
- Modélisation et simulation des systèmes urbains.

- **1 et 2 juillet** : Une exposition d'environ quatre-vingt stands regroupera les principaux acteurs du HPC : constructeurs et éditeurs, fournisseurs et intégrateurs de solutions matérielles, logicielles et de services, universités et laboratoires de recherche, etc.



Un forum pour échanger et pour informer les entreprises de l'intérêt du HPC et des possibilités d'y accéder.

démarche à l'échelle du vieux continent, même si celui-ci a déjà pris du retard comme le souligne Jean Gonnord, vice-président de Teratec : « L'Europe pèse 30 % du marché mondial en tant qu'utilisateur. Pourtant, elle ne fournit que 4 % des ressources, tandis que les supercalculateurs fabriqués

aux États-Unis accaparent 95 % du marché de l'Union ».

Reste que l'Europe représente une chance pour les entreprises françaises de prendre place dans la course mondiale du HPC. Le calcul numérique est un formidable accélérateur pour l'innovation notamment et la compé-

titivité globale en matière de recherche, de développement de la médecine, des énergies renouvelables, de la prédiction climatique, de l'industrie, de la chimie... et, d'une manière générale, pour relancer la croissance. « Cette prise de conscience à l'échelle supra nationale et les actions qui en découlent, notamment sous l'égide du 8^e PCRD : Horizon 2020, constituent les meilleures chances pour que notre pays puisse espérer rattraper ses concurrents » conclut Jean Gonnord.

C'est pourquoi en 2012 a été créée la Plateforme Technologique Européenne (ETP) pour le HPC. L'ETP4HPC regroupe les principaux acteurs du calcul intensif : Bull, Intel, IBM, STMicroelectronics, NVidia... des centres de recherche comme Inria, CEA ou Fraunhofer et des PME tels Caps Entreprises, Distene ou Enginsoft. Son objectif est simple : favoriser l'essor du HPC en Europe.

Pour cela, elle vient de définir les technologies prioritaires à développer au niveau de l'Europe pour aboutir à un écosystème solide en matière de HPC. Quatre axes stratégiques ont été identifiés dans ce SRA (*Strategic Research Agenda*) : les technologies pour atteindre l'Exascale, les technologies de base du HPC, le développement des usages, enfin la création d'un environnement global favorisant la réalisation des objectifs, notamment en terme de formation, et d'incitation à l'innovation et à la création d'entreprises.

Évidemment ces efforts de recherche nécessitent un financement. A travers des contrats Public/Privé (CPPP), l'Europe devrait investir chaque année entre 100 et 150 M € dans le cadre du programme Horizon 2020. Reste à savoir si ce montant sera suffisant face aux investissements des pays concurrents. ■

HPC : un facteur de compétitivité

La France s'est positionnée de longue date comme l'un des leaders européens en matière du calcul haute performance et de la simulation numérique. La course mondiale aux supercalculateurs les plus puissants est d'abord un enjeu d'innovation. Le plan « Supercalculateurs » lancé en 2013 par le gouvernement pour relancer la dynamique d'une France industrielle compétitive est présidé par Gérard Roucairol, Président de l'Académie des technologies et Président de l'association Teratec.

Cad-magazine : Quelles sont les actions retenues pour ce plan Supercalculateurs ?

Gérard Roucairol : Le plan se déploie en quatre volets. Le premier d'entre eux est la maîtrise des technologies propres au HPC en vue d'atteindre l'exaflopique en 2020. La France est l'un des rares pays dans le monde à disposer d'acteurs nationaux qui couvrent toute la chaîne de valeur de la simulation numérique. Il s'agit donc de s'appuyer sur cette force pour nous positionner aux côtés des leaders



Gérard Roucairol, Président de l'Académie des technologies et Président de l'association Teratec

que sont les USA, le Japon et la Chine. Cela signifie progresser sur tous les domaines : matériels, logiciels, programmation massive parallèle, applications, etc. Deuxième axe, la diversification des usages. Nous devons devenir moteur dans six grands domaines du HPC : la santé, la ville, les matériaux, les grands systèmes, le multimédia et le végétal. Cela passera notamment par la création de laboratoires communs, et le soutien de plateformes logicielles

métier par exemple. Troisième volet, la dissémination du HPC vers les PME et ETI françaises. Pour cela, nous devons croiser les initiatives des régions et des filières industrielles, avec la R&D et des services métier proposés par des entreprises commerciales. Enfin, la formation initiale et continue est un axe majeur pour progresser. Nous allons travailler avec l'Éducation Nationale pour renforcer les filières informatiques et modélisation mathématique.

Cad-magazine : Quels sont les moyens qui seront mis en œuvre ?

Gérard Roucairol : Tout n'est pas encore chiffré et il faudra nécessairement faire des arbitrages, car nous ne pourrions progresser sur tous les fronts à la fois. On peut estimer un budget minimal de l'ordre de 250 millions d'euros. Ces fonds seront majoritairement amenés par les entreprises qui s'inscriront dans ces projets, le reste étant à la charge de l'État à travers ses différents dispositifs d'aides.

Cad-magazine : Qui sont les acteurs de ce plan ?

Gérard Roucairol : On peut citer Vinci, Michelin, Air Liquide, Safran, L'Oréal, Ubisoft ou encore Airbus et bien entendu Bull pour les grands comptes. D'autres acteurs de plus petite taille comme Caps Entreprise ou des regroupements professionnels comme PFA (Plateforme Automobile) sont parties prenantes du plan. Notons également que certaines entreprises étrangères devraient participer.

Cad-magazine : Y a-t-il une roadmap, des objectifs précis et des critères d'évaluation d'un plan réussi ?

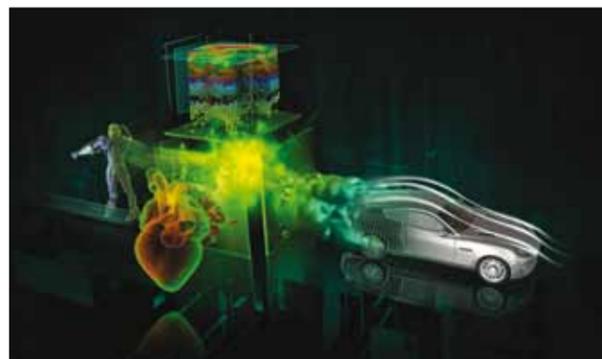
Gérard Roucairol : Bien entendu, il y a un agenda précis avec des rendez-vous réguliers. Mais, il s'agit de projets industriels impliquant de nombreux acteurs, avec des enjeux forts liés à la compétitivité nationale. Ce qui me contraint à une certaine discrétion sur ce type de détail. D'une manière générale, un bilan sera établi d'ici trois ans. ■

HPC et Cloud peuvent-ils faire bon ménage ?

Malgré d'évidentes similitudes en termes de besoins, le HPC et le Cloud conservent des différences fondamentales. Il existe cependant certains « use cases » pour lesquels le croisement des deux solutions se révèle utile pour des clients ou des services providers.

Les promesses du HPC

Inutile de revenir dessus, le calcul numérique est désormais un enjeu stratégique pour nombre de secteurs d'activité économique. Dans les industries qu'il sert, le HPC est au cœur du business, pour fabriquer de nouveaux produits, les rendre plus performants, plus économiques à produire, accélérer leur industrialisation, etc.



Dans quasiment tous les domaines d'activité économique, le calcul numérique s'est rendu indispensable. (Doc. Nvidia)

Pour tenir ces promesses, l'objectif est de poser un même morceau de code pour jouer une série d'opérations identiques sur un grand nombre de processeurs qui traiteront chacun un jeu de données spécifique. Les résultats de

l'ensemble des opérations identiques, sur des données distinctes, réparties sur plusieurs nœuds, sont ensuite consolidés.

Dans cette perspective, toute puissance de calcul supplémentaire apporte un bénéfice immédiat au métier de l'entreprise. Cela a donné lieu à une course effrénée à la puissance à travers notamment le développement d'algorithmes paral-

lèles, l'adoption d'architectures matérielles optimisées en fonction des besoins de performances des applications cibles, et la mise au point et la maintenance de suites logicielles et systèmes optimisées finement pour le matériel.

Les contraintes budgétaires et le besoin d'agilité dans la réponse aux besoins métiers s'appliquant à ce milieu comme aux autres, les centres de calculs ont adopté en parallèle un mode de fonctionnement de type « centre de services partagés ». Ce modèle a d'ailleurs sans doute inspiré les services de Cloud actuellement disponibles.

Alors, fusion Cloud et HPC ?

Le Cloud pourrait-il répondre au besoin de puissance de calcul des entreprises à travers le « débordement » sur une infrastructure HPC extérieure ? Pas si simple malheureusement, car le HPC présente un certain nombre de contraintes :

- Parce que l'adhérence fine entre les couches logicielles et le matériel est un enjeu dans le monde du calcul, l'utilisation d'une couche de virtualisation pose un problème, particulièrement en termes d'impact sur les performances.
- Le pilotage repose sur la visibilité fine de la disponibilité des ressources matérielles.

- Les temps de transferts de données potentiellement volumineuses entre différents clusters de calcul peuvent s'avérer prohibitifs.
- La sécurité est un point majeur étant donné les enjeux de propriété intellectuelle.

Alors on y va ou pas ? Il faut d'abord tordre le cou à la question de fond : « En définitive la virtualisation c'est possible pour le HPC ? » Dans la mesure où des expérimentations ont déjà permis de montrer que l'impact de l'usage d'un hyperviseur sur les performances, pouvait, suivant l'application cible, se révéler inférieur à 5 %, nous pouvons sérieusement l'envisager.

Avec quels bénéfices ?

L'adoption par les centres de calcul des technologies Cloud et en particulier de la virtualisation peut avoir deux avantages :

- Le « time to market ». L'usage des technologies de virtualisation qui sont au cœur du Cloud Computing introduisent une couche d'abstraction entre le matériel et le logiciel permettant ainsi de déployer de nouvelles applications avec moins de contraintes, donc plus rapidement.
- Le ROI. La capacité des Clouds les plus avancés à automatiser le déploiement de nouveaux services sur des architectures virtuelles aussi bien que physiques, ainsi que leur mode de gestion par pools de ressources physiques, permet plus aisément le



Le HPC peut-il se marier avec les technologies de Cloud Computing ? Il semble que oui...

partage de nouvelles architectures matérielles. La rentabilisation des architectures matérielles est accélérée, par rapport à un monde non virtualisé où la mise en œuvre d'un cluster dédié pour deux ou trois applications qu'il faudrait porter sur un socle logiciel contraint présenterait un ROI probablement trop long.

Enfin n'oublions pas le déplacement du barycentre CAPEX vers OPEX. La mise en place de solutions de débordement vers le Cloud également exploitables pour des applications de type HPC, rend possible la location de puissance à la demande pour gérer les pics de charge.

Pour quelles applications ?

Parmi les cibles naturelles pour bénéficier des technologies Cloud, je pense aux applications de type « embarrassingly parallel » avec un couplage faible et des données en majorité locales à chaque calcul. Ces applications sont plus fréquentes qu'on ne le croit : simulation financière de type Monte-Carlo, conception de composants électroniques, simulation dans le domaine de la recherche géné-

rique, mécanique des fluides, thermiques, etc.

Les caractéristiques des algorithmes utilisés dans ces applications leur permettront d'être exécutées avec une excellente performance sur des nœuds de calculs avec des caractéristiques techniques similaires voire identiques à celles d'un host de virtualisation habituellement utilisé pour une infrastructure de type Cloud privé ou public. Seule contrainte spécifique aux infrastructures de type Cloud HPC pour cette typologie d'applications : il faudra respecter des critères d'allocation de puissance plus restrictifs que pour les applications « classiques ». Aujourd'hui, les outils de gestion en mode Cloud sont à même de prendre en compte ce type de contrainte en sachant manipuler le provisioning d'environnement jusque dans les couches basses.

Cas spécifique du débordement

Les besoins à prendre en compte pour mettre en place une infrastructure de type Cloud pour du débordement HPC sont :

- L'envoi des données souvent sensibles, quelques fois volumineuses mais parfois « vivantes » dans le Cloud, ainsi que le rapatriement des résultats.
- Comme vu plus haut, l'allocation des ressources physiques, notamment les cœurs de CPU, ainsi que la mémoire vive dans le cas des applicatifs ciblés, aux « consommateurs virtuels » (vCPU et vRAM), suivant des ratios qui prennent en compte les caractéristiques de performance des applications, critère clé dans le calcul.
- L'ordonnancement (scheduling) des lots de calcul (batches) en fonction de l'usage réel du matériel et non de l'usage observé au niveau virtuel, afin de garantir la stabilité du point 2 dans le temps.

Ce point doit pouvoir être réalisé de manière industrielle et prédictible. Il ne pourra être répété que sur la base d'une solution industrialisée. Ceci peut être géré par l'usage, classique dans les infrastructures Cloud, d'outils de gestion de la capacité de distribution des changements qui vont adapter les politiques d'allocation de ressources des hyperviseurs. Un autre mécanisme complémentaire peut être l'usage des notions d'affinité et anti-affinité pris en compte par l'orchestrateur interne du composant Nova (Compute Service) d'OpenStack pour le placement des VMs.

A la question « Le HPC et le Cloud peuvent-ils faire bon ménage ? », la réponse est décidément oui. ■

Article réalisé en collaboration avec le site cloud-experience.fr