



JEAN-PHILIPPE NOMINÉ (83) *direction des Analyses stratégiques, CEA*

LA PUISSANCE DE CALCUL: UN MARIAGE ENTRE TECHNOLOGIES ET APPLICATIONS

Les besoins de la recherche et de l'industrie en matière de calcul intensif et de traitements de masses de données croissent à une telle vitesse que seuls des centres dédiés peuvent y répondre. Ces centres sont avant tout des centres de compétences où se mêlent grands équipements, méthodes de développement logiciel et support aux utilisateurs.

LES APPLICATIONS propres au monde des sciences et de l'industrie sont en moyenne plus intenses en calcul et traitements que les tâches transactionnelles au profit du Web, de la messagerie, des réseaux sociaux ou encore de maintes tâches d'informatique d'entreprise. En effet, la résolution des problèmes scientifiques concernés – tant dans la recherche académique que dans l'industrie – nécessite des capacités de calcul massivement parallèles et fortement couplées, avec gros débits et faibles latences de communication entre processeurs et serveurs. Cela disqualifie les solutions très distribuées, au profit

de supercalculateurs « localisés » et à réseau interne spécifique – en permanence au-delà des capacités de réseaux longue distance ou classiques.

DES ÉCURIES DE FORMULE 1

Beaucoup des technologies de base (processeurs, mémoires) sont communes entre supercalculateurs et serveurs plus génériques, même si le calcul haute performance (en anglais HPC, pour *High Performance Computing*) utilise plutôt le haut de gamme des composants produits par ailleurs en masse. Ce sont donc surtout la densité et l'intégration plus poussées qui caractérisent les architectures en

« Les plus grands centres de calcul scientifique actuels voient leur capacité mesurée en dizaines de pétaflops »

REPÈRES

Les centres de calcul dont il est question ici perpétuent le gigantisme des ordinateurs à tube des années 1940, occupant de grandes salles, en ayant bien entendu bénéficié de la miniaturisation de leurs composants intimes. S'ils sont parfois moins gigantesques que les plus grands centres informatiques des géants du Web, ou de la NSA américaine, ils sont plus « denses » : dans quelques centaines ou milliers de mètres carrés, et quelques mégawatts – parfois jusqu'à 20-30 – ils concentrent unités de calcul et de stockage.



Centre de traitement CEA.



Les équipes de support et d'exploitation peuvent regrouper des dizaines de spécialistes de nombreux sujets – une sorte d'analogie à une équipe technique d'écurie de Formule 1.

clusters utilisées et leur environnement local, qui comporte aussi des systèmes de refroidissement adaptés, et des systèmes de stockage de masse. Les plus grands centres de calcul scientifique actuels voient leur capacité mesurée en dizaines de pétaflops (calcul) et de pétaoctets¹ (stockage) – voir les statistiques sur www.top500.org. Les systèmes logiciels de ces machines sont adaptés à l'administration de masse, à la gestion de ressources partagées par des centaines d'utilisateurs distants, à la programmation parallèle massive des applications. Enfin, les équipes de support et d'exploitation peuvent regrouper des dizaines de spécialistes de nombreux sujets – une sorte d'analogie à une équipe technique d'écurie de Formule 1.

DES CENTRES DE COMPÉTENCES PLURIDISCIPLINAIRES

Penchons-nous sur les vertus et vocations de ces infrastructures de calcul et de traitement de données. Ces infrastructures sont privées (par exemple, propriété de grands industriels ou de groupements d'intérêt), ou publiques (au service de la recherche publique) et de différentes

échelles, croissantes : régionale, nationale, internationale. Les centres de calcul sont plus ou moins généralistes (privés, ils seront souvent plus spécialisés, mais des communautés de recherche peuvent disposer de centres qui leur sont réservés, comme la fusion pour l'énergie). Mais surtout, ce sont des centres de compétences : à la fois sur les technologies en perpétuelle évolution, sur les méthodes de développement logiciel, sur le support aux utilisateurs – en plus des missions de base de bonne exploitation des machines et de services de base associés.

« La France est ainsi l'un des rares pays à conserver une autonomie de développement de supercalculateurs »

LE COMPLEXE DE CALCUL DU CEA – TERA, TGCC, CCRT

À ce jour, le CEA exploite à Bruyères-le-Châtel dans l'Essonne deux des plus grands centres de calcul d'Europe, Tera, installation classifiée pour la défense, le TGCC, installation dédiée aux moyens et services pour la recherche et l'industrie – via 2 autres machines. La conjonction sur un même site de 3 machines « pétaflopiques » exploitées par les mêmes équipes CEA sur ces 3 grands axes d'usage du HPC, formant ce complexe de calcul, est quasiment unique au monde ; il en résulte des économies d'échelle et un partage extrêmement vertueux des compétences au profit de toutes les catégories d'utilisateurs. Le CCRT (partie industrie) en particulier est un montage très original de partage de machine avec plus de quinze industriels partenaires/utilisateurs. Voir www-hpc.cea.fr

UN PATRIMOINE À VALORISER

Cette notion de centre de compétences est en fait cruciale, elle doit être une ambition en soi ; il s'agit de faire fructifier le patrimoine inestimable de savoir-faire résultant de cette position privilégiée d'observatoire global : lieu de vie des grandes applications de calcul et traitement de données, permettant des optimisations plus globales et communes, des anticipations technologiques plus mûries et au final de meilleurs services hautement partagés, invoquant parfois des compétences très pointues et rares en traitement de données, administration de masse, algorithmique avancée, etc. De plus, les grands centres de calcul peuvent avoir une relation privilégiée avec les grands fournisseurs de technologie (composants ou systèmes) : accès et test précoces pour mieux anticiper les

LA COURSE « À L'EXASCALE » À TRAVERS LE MONDE

USA, Japon sont les grandes nations historiques du HPC et de leurs technologies. La Chine s'y investit désormais avec une volonté et des moyens considérables, en quête de souveraineté et d'indépendance, voire de domination sur le secteur. Tous ont des programmes « exascale », soit mille fois le pétaflops atteint vers 2010, ambitieux, clairement cadrés aux niveaux gouvernementaux. La Commission européenne est consciente de l'enjeu de la maîtrise du HPC comme fer de lance de nombreuses évolutions informatiques, et de l'importance du calcul intensif et traitement de données massives pour la compétitivité industrielle et scientifique. Un programme HPC global a donc été lancé au sein du programme cadre Horizon 2020. La France est le seul pays européen à avoir son propre programme d'ambition globale – technologies + infrastructures + applications – qu'il est impératif de combiner et d'amplifier au mieux avec les efforts européens, afin d'atteindre une masse critique de moyens de R & D, de compétences, de marché.



© EYETRONIC

La Chine investit dans ce domaine avec une volonté et des moyens considérables.

évolutions; voire conception conjointe des systèmes de calcul futurs: le CEA fait cela depuis quinze ans avec Bull (Atos) avec le soutien d'un Plan d'investissement d'avenir. La France est ainsi l'un des rares pays au monde à vouloir et pouvoir développer et conserver une autonomie de développement de supercalculateurs.

COLOCALISER DONNÉES ET TRAITEMENT

La place croissante prise par la problématique des données impose de penser une juste (co)localisation des calculs et des données. On rapproche aujourd'hui plus aisément un calcul de l'entrepôt de données que l'on ne déplace les pétaoctets. Les grands calculs produisent souvent de grandes masses de données locales; à traiter, interpréter, parfois redistribuer au mieux ensuite par parties; des grandes masses de données d'origine quelconque (calcul mais aussi, de plus en plus, issues *a priori* de capteurs, réseaux sociaux, etc.)

peuvent bénéficier de la proximité de forte puissance de calcul afin de les traiter et exploiter. On observe notamment une forte et rapide évolution des besoins d'analyse de données (*data analytics, machine learning, deep learning*) dont certains algorithmes pourraient dépasser les besoins en puissance de calcul de ceux de la simulation numérique – pas vouée à disparaître pour autant! Les grandes infrastructures de calcul sont donc bien des lieux privilégiés où faire vivre et évoluer la relation des grands calculs aux grandes masses de données, par évolution des architectures et services offerts, sur la base de forte compétence historique de gestion de systèmes et de services mutualisés complexes.

« Le calcul haute performance n'est pas aisément soluble dans du cloud totalement banalisé »

UN BESOIN DURABLE DE GRANDES INFRASTRUCTURES

Enfin, l'informatique en nuage (*cloud*) est perçue comme une tendance lourde pouvant reformuler potentiellement tous les usages de l'informatique; il s'agit en fait d'abord d'un mode d'accès à distance, et d'organisation plus « virtualisée » des services, que les grands centres de calcul pratiquent déjà très couramment. Le calcul haute performance n'est pas aisément soluble dans du *cloud* totalement banalisé; il requiert de toute façon les architectures techniquement adaptées décrites précédemment.

Les rendre accessibles en mode *cloud*, c'est surtout en augmenter la souplesse d'accès (« élasticité », qui peut aussi bénéficier à des utilisateurs plus occasionnels ou non institutionnels comme des PME ou des ETI).

En conclusion, calcul intensif et masses de données ont toujours coexisté en calcul scientifique. La croissance universelle des masses de données produites et/ou traitées dans des domaines de plus en plus variés ne fait que renforcer la valeur des grandes infrastructures de calcul capables de traiter les problèmes aux frontières. ■

TECHNOLOGIES ÉMERGENTES ET TECHNOLOGIES DE RUPTURE

Les recherches sont actives sur des voies alternatives aux circuits actuels (CMOS sur silicium) ou en rupture – ordinateur quantique, à ADN, neuromorphique, pour citer les principales voies. On ne voit aucune de ces dernières options à même d'émerger rapidement sur le marché et à un stade industriel, ni être « généralistes » dans leurs usages potentiels. Il importe néanmoins de suivre ces technologies, d'anticiper leur futur impact et notamment de penser leur programmation et exploitation efficaces. Les grands centres de calcul ont un rôle privilégié à jouer ici, surtout lorsqu'ils sont adossés et naturellement liés à des organismes technologiques et communautés de recherche au meilleur niveau mondial.

1. Péta = 10^{15} , soit un million de milliards ($\sim 2^{50}$).