

Contribution du GDR SoC² au rapport de conjoncture 2018 du CoNRS section 7

1. Les grandes avancées actuelles et les thèmes émergents/ enjeux de la discipline et quelle est sa place au sein de la société

La « *digitalisation* » de la société est sans aucun doute la quatrième révolution industrielle. Comme les précédentes, elle impacte l'ensemble de la société dans ses usages et ses attentes. De nouveaux défis sociétaux font leur apparition comme l'industrie du futur, l'e-santé, et les "*smart-cities*". Ces grands défis sociétaux ont un impact important sur la conception des systèmes embarqués (au sens électronique et informatique), supports indispensables de ce monde numérique.

La conception de ces systèmes est une thématique de recherche multidisciplinaire cruciale pour le déploiement de **l'électronique ambiante fortement contrainte par des exigences de temps de développement, de performances, de sûreté, de sécurité et d'efficacité énergétique**. Les nouveaux enjeux et nouvelles applications nécessitent aussi l'intégration, en plus du ou des cœurs de calcul, de capteurs et actionneurs innovants (comme par exemple pour l'e-santé avec des capteurs biocompatibles). On parle alors de **systèmes cyber-physiques** (CPS : *Cyber-Physical Systems*).

Enfin, cette numérisation s'appuie sur le nomadisme permettant à chacun d'accéder à un espace numérique ambiant dans lequel les accès aux informations et services sont transparents et dont le contrôle est complètement décentralisé. Ceci requiert alors des développements en communication numérique, afin de gérer les canaux de communication avec un nombre d'objets communicants dépassant plusieurs milliards. Cet ensemble distribué devra de fait s'adapter et évoluer de façon plus ou moins autonome, cette propriété s'appuyant alors sur l'intelligence artificielle. La mise en réseau de nos données pose alors la question de la sécurité et de la fiabilité de ce monde numérique distribué.

Du point de vue des systèmes matériels eux-mêmes, les principaux challenges actuels s'articulent autour de quatre préoccupations essentielles :

- l'autonomie, adaptabilité et l'**efficacité énergétique**,
- la **mise à l'échelle des performances**,
- la **prédictibilité temporelle**,
- la **résilience et la sécurité**.

1.1. Edge computing

La prise de conscience notable de la question du développement durable dans le contexte de l'informatique dans le nuage, reposant sur des centres de calcul présentant un impact carbone significatif, motive de nombreux travaux visant à augmenter l'efficacité énergétique des systèmes de calcul. Cette tendance se décline à la fois s'agissant des systèmes de calcul hautes performances (HPC) mais aussi des nœuds légers tels que prônés par le "**edge computing**". Dans de nombreux domaines d'application, la quantité de données produites ne cesse de croître, ce qui pose alors des

questionnements sur les traitements de celles-ci, non seulement sur les types de traitement à réaliser, mais également sur la localisation de ces traitements. Si un traitement centralisé peut avoir du sens lorsque le volume de données est maîtrisé, il n'est plus tenable lorsque la quantité de données produites par les capteurs distribués augmente fortement (cas de l'IoT). Il faut alors envisager des **traitements au plus proche de la source des données** pour ne faire remonter vers un point centralisé que des informations sémantiques utiles à une prise de décision ou à un traitement global.

C'est dans ce contexte que le concept du "**edge computing**" et du "**near-sensor computing**" se développe et tente de répondre à cette problématique. Elle n'est pas nouvelle pour certains domaines : on citera par exemple le cas des prétraitements d'images réalisés par des rétines afin de soulager le processeur de certains travaux fastidieux et consommateurs de puissance de calcul. Toutefois, les challenges liés à cette distribution des traitements sont nombreux puisqu'il s'agit de trouver le bon compromis entre la localisation des traitements et le coût des communications. Une autre approche pour améliorer l'efficacité énergétique s'appuie sur les évolutions technologiques, et notamment l'apparition des mémoires non-volatiles. Ainsi, basé sur l'utilisation de ces technologies et sur la création de systèmes de récolte de l'énergie ambiante (*harvesting*) il est possible de créer des objets autonomes en énergie, voire ne s'allumant que lorsque l'on en a besoin – il s'agit des concepts de **normally-off computing**, ou de **transient computing**. Ces concepts nécessitent, hormis de nouvelles approches architecturale et technologique, des travaux sur les méthodologies de gestion de ces architectures et notamment des systèmes d'exploitation adaptés prenant en compte cette intermittence de calcul.

1.2. Intelligence artificielle

Un des premiers domaines à adresser est celui du "**Machine Learning**" dont non seulement l'entraînement mais également l'inférence sont particulièrement exigeants en terme calculatoire. L'effervescence des recherches autour de l'intelligence artificielle et l'apparition de nombreux réseaux de neurones présentant des caractéristiques et des objectifs différents auront un impact important sur la conception des systèmes cyber-physiques. Ainsi, les réseaux de neurones profonds, les réseaux convolutifs ou les réseaux génératifs, pour ne citer que ceux-là, nécessitent des outils, méthodes et architectures spécifiques pour une implémentation efficace. Les travaux concernent souvent la proposition de solutions innovantes, reposant sur des solutions matérielles (*in-memory computing*, architectures *manycores*) ou logicielles à travers de nouvelles approches de gestion en ligne (*runtime*) ou de compilation. Les supports de calcul sont systématiquement parallèles et d'un niveau de complexité important. Cela rend difficile la question de la prédictibilité temporelle, et donc de leur exploitation efficace dans les systèmes critiques. Des travaux sont donc menés à différents niveaux (compilation, microarchitecture, analyse statique, ordonnancement des calculs et des accès mémoires) pour maîtriser les performances au pire cas ou en temps moyen. Enfin, on note un engouement particulier s'agissant de l'exploitation de nouvelles technologies de fabrication (mémoires non volatiles notamment) dans le contexte de la réalisation de systèmes matériels, généralistes (insertion de mémoires non volatiles dans la hiérarchie mémoire d'ordinateurs) ou spécifiques (architectures neuromorphiques par exemple).

1.3. Sécurité

D'un point de vue plus global, l'ensemble des travaux ci-dessus devront prendre en compte des contraintes de sécurité et de fiabilité. La sécurité est un enjeu majeur avec de nouvelles problématiques, notamment pour l'IoT. Ainsi tous les éléments de l'espace numérique devront être protégés contre des menaces extérieures afin de protéger les données et les personnes. Ainsi, il faudra notamment envisager la réalisation efficace des nouveaux schémas de chiffrement. Dans le cadre des systèmes cyber-physiques, il conviendra de réaliser des implémentations matérielles légères, et même

ultralégères, permettant de satisfaire aux contraintes de consommation et performances des systèmes embarqués. Cette propriété de sécurité nécessite une approche holistique. Ainsi, à titre d'exemple, les travaux de recherche concernent l'identification et authentification de systèmes en déployant des solutions matérielles de confiance, l'intégrité et la confidentialité des données par des solutions compatibles avec les besoins énergétiques actuels, l'analyse sécuritaire (attaques et contremesures). Des travaux amonts sont à considérer notamment sur la cryptographie homomorphe, qui permet une sécurisation forte des données mais dont la complexité actuelle n'est pas compatible avec une approche embarquée. Il sera également nécessaire de concevoir des nœuds de calculs intrinsèquement sécurisés (concept de "*secured by design*"). Finalement, des outils d'évaluation ou de vérification formelle de la sécurité et de la sûreté seront requis.

1.4. Nouvelles technologies, nouveaux paradigmes

Les évolutions technologiques (photonique, spintronique, quantique, etc.) auront dans un futur proche un impact important sur la conception des systèmes cyber-physiques et il convient dès lors de s'intéresser à ces avancées technologiques afin d'anticiper de futurs changements de paradigme de calcul. En effet il est important de ne pas dissocier les paradigmes de calcul avec l'évolution actuelle des technologies, l'un étant indissociable de l'autre. Plusieurs axes de recherche sont à développer autour de la cryptographie post-quantique, du calcul quantique, d'architectures neuro-inspirées, d'architectures de processeurs résilients, la prise en compte de nouvelles mémoires pour les systèmes embarqués etc. Il est donc indispensable de favoriser les liens entre technologies et nouveaux paradigmes de calcul pour relever les défis des systèmes ambiants de demain.

2. Place de la France dans le contexte de la recherche mondiale

L'avènement des objets connectés et des approches « matériel libre » (à l'instar du jeu d'instructions libre de processeur RISC-V) ouvrent de grandes opportunités pour la recherche en France et la conception des objets connectés en particulier. En effet, la France possède une industrie et une recherche riches et foisonnantes avec des leaders mondiaux (Thales, ST, Airbus, Valeo, etc.) qui portent notamment la création d'une filière stratégique dédiée à l'électronique. Ainsi le projet européen ECSEL CPL ou le projet Nano2022 visent à asseoir à l'international, les avancées Françaises en matière d'électronique et d'architecture, mais aussi de conception des systèmes cyber-physiques.

Les recherches et les laboratoires impliqués dans cette thématique sont organisés depuis très longtemps au sein du GdR SoC². Le domaine couvert par le GdR est traditionnellement considéré comme un domaine fortement interdisciplinaire entre les aspects matériels et logiciels, interdisciplinarité indispensable à la conception des systèmes cyber-physiques. Le GdR SoC² occupe ainsi une position originale à l'interface entre les communautés électronique, physique et informatique, et donc à l'interface entre les Instituts INSIS, qui est son rattachement principal, et INS2I du CNRS. Au CoNRS, ses membres sont principalement rattachés aux sections 7 et 8.