

# Contribution du GDR SoC<sup>2</sup> au rapport de conjoncture 2018 du CoNRS section 8

## **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 majeur 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 – en tout premier lieu l'efficacité énergétique, de sûreté et de sécurité**. 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 et sécurisé. 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'efficacité énergétique**, l'autonomie et adaptabilité,
- **l'amélioration des performances** pour suivre les besoins toujours croissants des applications,
- la construction de support matériel pour **l'intelligence artificielle**,
- la **résilience et la sécurité**.

### **1.1. Edge computing**

La volonté de réduire l'impact énergétique et carbone des centres de calcul conduit à repenser la localisation des traitements et à réduire la quantité de données transférées depuis le capteur jusque dans le cloud. Dans ce contexte, le paradigme du "edge-computing" (calcul proche capteur) répond à cette problématique puisqu'il propose de placer des traitements de données au plus proche du capteur afin d'extraire des informations pertinentes qui seront alors renvoyées vers les centres de calcul situés dans le cloud. Les challenges scientifiques de ce domaine concernent alors tous les niveaux de conception des capteurs et du calcul associé.

On citera par exemple l'utilisation de technologies qualifiées émergentes telles que les mémoires non-volatiles, qui atteignent pour certaines un niveau de maturité qui autorise leur exploitation dans divers types d'architectures de calculs. Ces composants mémoire permettant de conserver des données ou des états de calcul, tout en étant majoritairement en mode ultra faible consommation ("normally-off" ou "transient" computing). De par leur caractère non-volatile, le système de calcul est capable de se réveiller et de se mettre en veille de manière quasi-instantanée et transparente, pour une meilleure efficacité énergétique. Les enjeux consistent ainsi en la réduction du temps et de l'énergie d'écriture et de lecture, et en l'endurance (nombre de cycles d'écriture).

On assiste également à l'émergence de solutions technologiques hybrides mémoire centrale / mémoire de stockage vouées à prendre place dans des architectures de calculateurs conventionnels (calcul intensif, datacenter) alors que plusieurs démonstrateurs technologiques soulignent également l'intérêt de l'hybridation technologique – la coexistence de ces mémoires émergentes et de mémoires conventionnelles au sein même des processeurs (mémoires caches, scratchpad memories).

De nouveaux paradigmes de calcul s'appuyant sur le calcul au plus proche des mémoires (logic-in-memory) demandent également une intégration plus intime d'éléments mémoire avec les composants de calcul. Les approches d'intégration hétérogène tri-dimensionnelle pour les composants mémoires intégrés dans le Back-End-Of-Line (OxRAM, MRAM) et les performances des sélecteurs (vitesse, compacité, surcoût énergétique) sont aujourd'hui largement admises mais pour lesquelles les solutions d'avenir ne sont pas encore disponibles ; d'autres approches plus prospectives telles que l'intégration de matériaux ferroélectriques dans le Front-End-Of-Line pourront autoriser des stratégies de calcul plus agressives en termes de budget performances / consommation, ainsi que des supports matériels pour l'intelligence artificielle.

On peut également citer les techniques de compression de données, dont l'objectif est d'extraire l'information utile en vue de réduire les flux de données transférés dans le réseau. L'interface numérique (conversion analogique-numérique) est également un point critique dans un contexte faible consommation, avec une exigence de grande sensibilité. On notera alors des approches avec des architectures de systèmes incluant le capteur au sein du dispositif de conversion ainsi que d'autres proposant une compression des données par des schémas d'échantillonnage non uniforme. Les techniques de distribution de calculs sont également critiques dans un contexte où les capteurs co-opèrent pour fusionner de l'information ou simplement pour la remonter au niveau supérieur.

Enfin, l'alimentation de ces capteurs est un enjeu important, et les techniques de récupération d'énergie sont essentielles pour assurer une durée de vie et / ou de fonctionnement suffisante. Finalement, l'ensemble de 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 l'intermittence des calculs, la veille profonde, la récupération d'énergie et sa gestion au plus près des besoins.

## **1.2. Intelligence artificielle (IA)**

L'intelligence artificielle, et tout particulièrement l'apprentissage automatique / statistique, est au cœur des révolutions du numérique. La complexité calculatoire relative non seulement à la phase d'entraînement mais également à l'exploitation (inférence) implique l'utilisation de ressources matérielles de calcul spécialisées (accélérateurs), avec un impact énergétique très significatif crucial aux préoccupations d'efficacité énergétique et de développement durable, notamment dans les datacenters et les systèmes relevant du « Edge Computing ». Les avancées dans le domaine des technologies émergentes combinées à l'émergence d'architectures de calculs neuromorphiques exploitant ces technologies permettent aujourd'hui d'envisager l'exploitation des technologies relevant de l'intelligence artificielle à divers niveaux et dans des enveloppes énergétiques largement réduites en comparaison des solutions actuelles.

### **1.3. Sécurité**

La sécurité des systèmes-sur-puce complexes et hétérogènes qui équipent de plus en plus de systèmes électroniques connectés est un enjeu majeur aux nombreux verrous : efficacité des attaques physiques avec la réduction de dimensions des technologies et l'augmentation de la complexité, mise en évidence de nouveaux chemins d'attaques internes du matériel au logiciel et vice versa, sécurité vis-à-vis des moyens de tests et de correction d'erreurs, ... L'implantation matérielle efficace (consommation énergétique, sécurité, performance, surface) de nouveaux schémas de chiffrement, tels que le chiffrement homomorphe et le chiffrement post-quantique, est une problématique actuelle importante. Enfin dans le domaine des attaques par canaux cachés (analyse de consommation de puissance et / ou électromagnétique) des implémentations cryptographiques matérielles et logicielles, l'utilisation récente des techniques d'apprentissages automatiques (machine learning, deep learning) a donné de très bons résultats en termes d'efficacité même sur des implémentations sécurisées. Ces avancées ouvrent un nouveau champ de recherche sur l'exploitation de l'IA pour la sécurité des systèmes matériels.

### **1.4. Nouvelles technologies, nouveaux paradigmes**

Les technologies émergentes micro-nanoélectroniques permettent de réaliser des fonctions de calcul (nanofils, nanotubes, SET, etc.), de mémorisation (MRAM, ReRAM, etc.) ou de communication (nanophotonique, spin, terahertz, optique visible, etc.), dont l'impact sur les futurs systèmes est attendu en termes d'efficacité énergétique, de puissance de calcul et de robustesse (fiabilité et sécurité) des SoC. Bien entendu, cette attente s'avère importante dans le cadre des paradigmes conventionnels de calcul. De façon plus prospective, ces mêmes évolutions technologiques et d'autres comme la photonique sur silicium, la spintronique ou le calcul quantique ouvrent également la voie à de nouvelles architectures basées sur des paradigmes plus en rupture.

Le déploiement de ces nouveaux paradigmes de calcul comme la logique multi niveaux, le calcul approximatif, asynchrone, neuro-inspiré, etc. s'anticipe en tirant profit des propriétés innovantes des technologies émergentes. Il s'agit alors de mener une réflexion sur les paradigmes en rupture qui entraînent des convergences calcul/mémoire, calcul/communication et mémoire/communication. Outre le recensement de ces technologies émergentes, leur degré de maturité régulièrement mis à jour est une des missions du GdR SOC2 tout comme l'anticipation de leur passage à l'échelle (celle du circuit) qui passe par la conception des modèles compacts associés et l'évaluation de leur niveau de fiabilité.

### **1.5. Communication haut-débit**

De nos jours, les systèmes informatiques en réseau sont de plus en plus présents dans nos vies quotidiennes et les systèmes électroniques ont de plus en plus d'interactions avec le monde physique, de sorte que le monde numérique et le monde physique fusionnent pour évoluer vers des systèmes cyber physiques. Cette transformation vise à accroître la capacité d'adaptation, l'autonomie, l'efficacité, la fonctionnalité, la fiabilité et la sécurité des futurs systèmes. Le verrou est de développer des éléments matériels permettant de relier à faible coût le système d'objets connectés, d'un côté au monde physique via des capteurs, et de l'autre côté au monde numérique via des communications sans fil. La prise en compte et la gestion de la consommation orientent la conception des circuits analogiques / RF / hétérogènes vers plus de modularité et d'agilité et une plus grande intégration. Les interfaces de communication intelligentes conçues à cette fin utilisent une interaction accrue avec le monde physique (capteurs intelligents, récupération d'énergie), la connaissance du réseau, la puissance du calcul numérique (adaptation, correction) et la connaissance de l'usage auquel les données sont destinées.

Les émetteurs-récepteurs 5G devront faire preuve d'une grande flexibilité dans de nombreux aspects de conception, tels que la fréquence centrale, le filtrage et le facteur de bruit, afin de satisfaire aux spécifications tout en maintenant une faible consommation d'énergie. Les techniques de correction

numérique seront de plus en plus utilisées pour corriger les imperfections dans les circuits analogiques / mixtes et RF.

Pour réduire la consommation d'énergie, l'enjeu est de réinventer les parties communication et détection des capteurs. La partie détection consistera à n'acquérir que les données pertinentes en ajoutant de l'intelligence au capteur. Pour ce faire, il est possible de combiner des techniques de détection récentes (analogique-information et analogique-paramètre) avec un apprentissage automatique afin d'optimiser la consommation et de réduire le volume de données à transmettre. La partie communication visera à proposer de nouveaux récepteurs de réveil intelligents capables de détecter la présence du signal en plusieurs étapes afin de maximiser le temps passé en mode veille et de réduire ainsi la consommation d'énergie.

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

La place de la France dans le contexte de la recherche mondiale relatif aux systèmes et objets connectés est tout à la fois très pertinent et tout à fait opportun. S'agissant du contexte national, notons l'importance du plan numérique français pour les systèmes électroniques d'une façon générale et en particulier de son écosystème micro/nanoélectronique qui a été largement mis en avant par le CNRS durant l'année écoulée via son "initiative microélectronique". Cet écosystème est "drainé" par deux grands industriels du secteur de classe mondiale, STMicroelectronics d'une part pour les moyens de fabrication et THALES pour l'ingénierie des systèmes. S'agissant du contexte européen, notons la prise de conscience de sa nécessaire souveraineté vis-à-vis de la puissance supposée de la convergence de l'Intelligence Artificielle (AI) et de l'Internet des objets (IOT), dont le concept est désormais appelé AIOT. Dans ce double contexte, national et européen, deux grandes exigences sont mises en avant de façon récurrente. D'une part l'efficacité énergétique, qui permet un déploiement de calcul et/ou des transferts de données à un coût énergétique minimal. Et d'autre part une co-optimisation hardware/software qui puisse répondre efficacement à l'application visée. Les paragraphes précédents ont montré combien, de par les objets qu'il manipule, le GDR SOC2 répond naturellement à ses deux exigences. En effet, la recherche de l'efficacité énergétique est une constante historique des systèmes et objets connectés sans laquelle l'alimentation en énergie s'avère un frein au déploiement d'une nouvelle application. De même, la co-optimisation hardware/software est une nécessité dont dépendent tous les systèmes embarqués et les "System On Chip", SOC en particulier. Ainsi donc, le GdR SOC2 est au cœur des problématiques de recherche de la quatrième révolution industrielle dont découleront les emplois de demain.