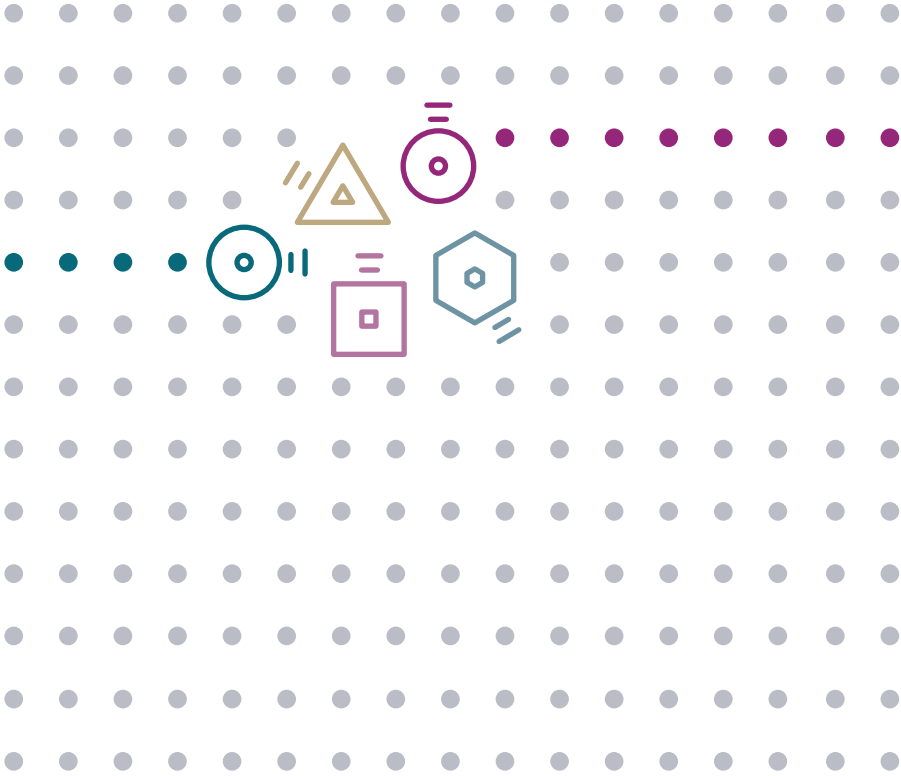
Internet des Objets

Défis sociétaux et domaines de recherche scientifique pour l'Internet des Objets (*IoT*)



Inria | LIVRE BLANC | N° 05

Ce document a été rédigé et coordonné par Emmanuel Baccelli.

***Contributions** – Merci aux personnes suivantes pour les contributions et les éclairages qu'elles ont apportés : N. Anciaux, K. Bhargavan, G. Casiez, F. Gandon, N. Georgantas, J.M. Gorce, S. Huot, E. Lank, A. Lebre, W. Mackay, K. Marquet, N. Mitton, E. Rutten, M. Tommasi, P. Vicat-Blanc, O. Sentieys, B. Salvy, M. Serrano, B. Smith, M. Vucinic, T. Watteyne.*

***Remerciements** – Que les personnes dont les noms suivent soient remerciées pour leurs commentaires éclairés : C. Adjih, F. Baccelli, C. Bormann, I. Chrisment, F. Cuny, F. Desprez, MA Enard, JF Gerbeau, P. Guitton, V. Issarny, P. Jacquet, L. Mé, V. Roca, H. Tschofenig, A. Viana..*

Date de Publication : novembre 2021

Sommaire

Résumé **03**

AVANT-PROPOS : Alice vit dans un cocon **06**
Le rêve d'Alice 06
Le cauchemar d'Alice 07
Libérer l'optimisme et atténuer le pessimisme associés à l'IoT 08

PARTIE 1 - L'internet des objets (IoT) **10**

1. IoT : Hier, aujourd'hui, demain **11**
1.1 Rapide historique de l'IoT 12
1.2 Les innovations permettant l'émergence de l'IoT 13
1.3 L'IoT aujourd'hui : le point de bascule 17
1.4 Quelles perspectives pour l'IoT ? 18

2. Les défis sociétaux de l'IoT **20**
2.1 Le cadre juridique : l'équilibre entre innovation sans permis et soucis éthiques 21
2.2 La confiance de l'opinion publique : la gagner et la conserver 22
2.3 Souveraineté 23
2.4 Normalisation 24
2.5 Éducation 25
2.6 Lutter contre le changement climatique et la raréfaction des ressources 26
2.7 Comment Inria contribue à répondre aux défis sociétaux de l'IoT ? 27

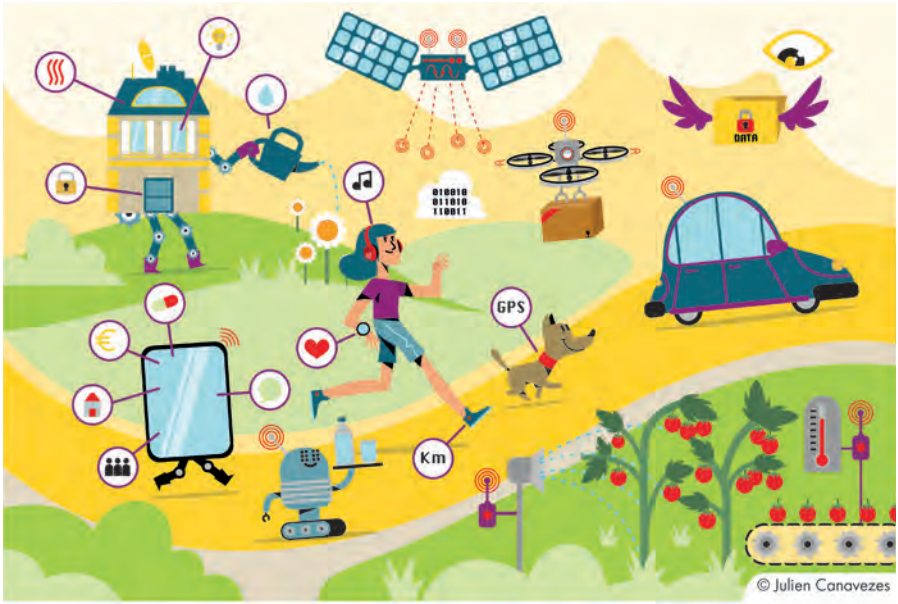
3. Les défis scientifiques et techniques de l'IoT **29**
3.1 Comment concilier sphère privée et IoT omniprésent ? 30
3.2 Comment renforcer la résilience, la sûreté et la sécurité de l'IoT ? 30
3.3 Comment associer apprentissage automatique et IoT ? 31
3.4 Comment étendre la connectivité de la boucle locale pour l'IoT ? 32
3.5 Comment repousser les limites des concepts réseaux de bout en bout pour l'IoT ? 33

3.6 De quelles interfaces Homme-machine l'IoT a-t-il besoin ? Et quelles interfaces découlent de l'IoT ? 33
3.7 Comment créer des passerelles entre IoT, contrôle et robotique ? 34
3.8 Comment concevoir des appareils IoT miniaturisés à l'échelle du millimètre ? 34
3.9 Comment tendre vers la neutralité pour l'empreinte de l'IoT sur les ressources naturelles ? 34
3.10 Comment Inria contribue à répondre aux défis scientifiques et techniques de l'IoT ? 35

PART 2 - Domaines de recherche de l'IoT **36**

Réseaux de communication pour l'IoT 37
Représentation de données pour l'IoT 44
Systèmes distribués pour le continuum *Cloud-Edge*-Objet 47
La cryptologie appliquée aux objets connectés « bas de gamme » 56
Traitement et confidentialité des données dans l'IoT 62
Sûreté, fiabilité et certification pour l'IoT 68
Interaction Homme-machine avec l'IoT 70
Contrôle avec l'IoT dans la boucle 74
La sécurité dans l'IoT 78
Architecture matérielle de faible puissance, programmation et compilation 89
Optimisation de l'empreinte ressources globale 97

CONCLUSION **102**



Résumé

De la même manière qu'Internet a profondément bouleversé notre société, l'Internet des Objets (*Internet of Things*, ou "IoT" en anglais) impactera tous les secteurs de l'activité humaine : notre habitat, nos véhicules, notre environnement de travail, nos usines, nos villes, notre agriculture, nos systèmes de santé... De même, tous les niveaux de la société (individus, entreprises, États) sont d'ores et déjà concernés, de l'urbain au rural, ainsi que la nature au-delà. Dès lors, comprendre les fondements et les enjeux de l'IoT apparaît crucial.
Ce document a en premier lieu pour but de :

- définir les contours de l'IoT, sa genèse, son actualité et ses perspectives ;
- identifier les principaux défis sociétaux, techniques et scientifiques relatifs à l'IoT.

Une forte accentuation, jusqu'à l'omniprésence de l'IoT, paraît inéluctable. L'IoT a en effet vocation à s'insérer dans les moindres aspects de la vie de tout un chacun, connectant tout (des milliards de nouvelles machines hétérogènes communicantes) et mesurant tout de nos agissements collectifs à l'échelle planétaire et au-delà, à nos plus infimes signaux physiologiques individuels, en temps réel. Cette vocation est à double tranchant : elle défie l'imagination pour le meilleur (automatisations, optimisations, fonctionnalités innovantes...)

comme pour le pire (surveillances, dépendances, cyberattaques…). L'IoT étant en perpétuelle évolution, de nouveaux défis sociétaux concernant la protection de la vie privée, la transparence, la sûreté et de nouvelles responsabilités civiles ou industrielles, commencent à apparaître.

L'IoT s'appuie sur un ensemble de plus en plus complexe de concepts et de technologies imbriqués et enfouis. Pour un acteur industriel, cette complexité grandissante rend plus difficile (voire illusoire) d'envisager seul une maîtrise fine, de bout en bout, des éléments constitutifs de l'IoT. Néanmoins, la culture générale de demain devra permettre d'en appréhender les fondements technologiques. Un défi pour l'enseignement est donc d'augmenter progressivement la sensibilisation à l'IoT, à la fois pour préserver la souveraineté et le libre arbitre des individus, et pour mieux amorcer les formations de nos scientifiques et nos techniciens. Un institut public de recherche tel qu'Inria peut contribuer à la fois à maîtriser et à expliquer les fondements technologiques de l'IoT, ainsi qu'à préserver la souveraineté en Europe.

L'IoT augmentera inévitablement la dépendance à certaines technologies enfouies. Ceci implique d'identifier les nouveaux risques, et d'élaborer de nouvelles stratégies pour tirer tous les bénéfices de l'IoT, tout en minimisant ces risques. Comme dans d'autres domaines où il faut chercher à préserver continuellement l'éthique sans pour autant entraver l'innovation, l'encadrement de l'IoT par la Loi est un effort à la fois nécessaire et ardu. Il semble toutefois clair que le niveau européen soit le niveau adéquat (comme le montre le RGPD par exemple) pour peser face aux géants industriels ou aux superpuissances. Par ailleurs, les normes technologiques ayant une influence grandissante sur notre société, il paraît indispensable de participer activement aux processus de normalisation de ces technologies. Les normes ouvertes notamment, ainsi que l'*open source* conçu comme « bien commun public », seront des moteurs de premier plan pour l'IoT tout comme ils l'ont été pour Internet.

Enfin, le défi environnemental auquel nous faisons face pourra être mieux capturé, et on l'espère, atténué, grâce à une utilisation massive de l'IoT. Il ne s'agit pas seulement de réduire le coût en ressources naturelles consommées par l'IoT (pour sa production, son déploiement, son entretien, et le recyclage). Il s'agit aussi d'être en mesure de pouvoir évaluer plus précisément, à l'échelle planétaire, le bénéfice net global de l'IoT sur l'environnement, en défalquant son coût environnemental des bénéfices attestés qu'il apporte, ce qui relève de la gageure actuellement.

L'impact grandissant de l'IoT souligne la nécessité de se maintenir à la pointe des développements technologiques et de la recherche qui le sous-tendent. Ce document a donc en second lieu pour but de :

- mettre en lumière la diversité des domaines de recherche sur lesquels s'appuie fondamentalement l'IoT ;
- passer en revue les problématiques de recherches actuelles et futures dans chacun de ces domaines.

Au gré du document, un certain nombre de liens sont établis avec les contributions d'Inria. Ces dernières sont de natures diverses (recherche fondamentale et appliquée, logiciel libre, incubation de startups...) et concernent la plupart des domaines de recherche sur lesquels s'appuie l'IoT.

AVANT-PROPOS : *Alice vit dans un cocon*

Le rêve d'Alice

La liberté compte beaucoup pour Alice. Si Alice le décide, ses appareils, vêtements et implants intelligents (combinant capteurs, actionneurs et communications locales sans fil) peuvent travailler ensemble. Alice peut aussi facilement connecter ses appareils à d'autres équipements intelligents qui se trouvent à proximité, ou à des solutions informatiques distantes qu'elle choisit sur le réseau, en fonction de ses besoins du moment. Globalement, le système constitue pour elle une forme de cocon personnel et cyberphysique qui « amortit » son expérience du réel, qu'elle soit chez elle, sur la route, ou au travail.

Pour se rendre sur son lieu de travail à l'usine, Alice utilise un véhicule autonome qui interagit avec les infrastructures de la ville intelligente pour emprunter automatiquement l'itinéraire le moins pollué en tenant compte de ses préférences, ou pour trouver une place de stationnement proche de sa destination. Les systèmes avancés de maintenance prédictive et de surveillance en temps réel de l'environnement que l'usine utilise garantissent la sécurité et la productivité du lieu de travail, tout en optimisant les consommations d'énergie. Alice ne quitte pas son cocon en rentrant chez elle : il personnalise son expérience cyberphysique, prenant en charge toutes ses interactions avec ses appareils. Et plus important que tout : elle garde le contrôle et peut faire confiance à un système dont le fonctionnement garanti et sécurisé préserve sa vie privée.

Alice utilise par exemple son cocon pour l'aider à prendre soin de sa santé avec un accompagnement prédictif. Elle peut bien sûr choisir de désactiver le système et de supprimer les données enregistrées à tout moment, à partir d'une interface simple mais complète. Elle peut facilement changer d'appareil ou décider quelles ressources informatiques connecter. Alice peutégalement consulter les principaux paramètres du système au travail, et réaliser un autodiagnostic de son état de santé, par exemple en recevant des alertes en cas de problème possible. Si elle le souhaite, Alice peut choisir de partager certaines de ses données de santé avec son médecin, provisoirement ou à long terme. Sur demande, son «cocon numérique» peut jouer un rôle plus actif dans ses traitements médicaux ou contribuer à prévenir de façon coordonnée la propagation de certains virus.

Alice bénéficie ainsi des meilleurs soins, si tel est son choix, à un coût optimal pour elle, son employeur et pour la société dans son ensemble. Sans compter qu'elle peut accéder à tous ces services, alors qu'elle vient tout juste de déménager à la campagne !

Le cauchemar d'Alice

Le taxi autonome d'Alice a percuté un arbre. La société de taxi expliquera plus tard que sa flotte de véhicules a connu des dysfonctionnements en raison de l'usurpation de signaux GPS. Alice n'a pas été gravement blessée heureusement, mais l'accident a endommagé l'un de ses précieux implants connectés. Il a fallu plus de temps que prévu à Alice pour rétablir le bon fonctionnement de son cocon numérique (et remplacer un appareil défectueux). Et comme si cela ne suffisait pas, elle s'est retrouvée contrainte de payer une forte « rançon » à des pirates qui ont exploité une faille de sécurité de son autre implant.

Alice doit maintenant réduire son train de vie. Sous la pression de sa compagnie d'assurance, dont l'unique objectif est de limiter les risques et de maximiser les bénéfices, elle n'a d'autre choix que d'accepter d'utiliser d'autres appareils surveillant ses signes vitaux, et de voir ses données confidentielles vendues à des tiers. Sa vie privée est... de moins en moins privée. La réalité surpassant douloureusement l'imagination d'Orwell, chaque geste d'Alice est maintenant scruté dans les moindres détails, en temps réel, tandis que différents actionneurs déforment subrepticement sa perception du réel.

Au travail, elle est constamment épiée par ses superviseurs, qui abusent des étonnantes capacités de surveillance et d'ingérence des déploiements de l'IoT dans l'usine, dispositifs qui sont à leur tour régulièrement victimes de cyberattaques, compromettant la productivité autant que la sécurité sur son lieu de travail. Dans sa sphère «privée», Alice est la proie du « capitalisme de surveillance ». Ses choix sont souvent influencés par des sociétés avides de profit qui exploitent son cocon cyberphysique.

S'appuyant sur une étude pilote, le gouvernement propose une nouvelle politique généralisant l'utilisation obligatoire des capteurs médicaux du cocon numérique, sacrifiant au vu et au su de tous l'intimité de ces données à l'objectif de réduction de la dette publique du pays. Cette réforme est débattue dans un contexte de forte suspicion concernant les récentes élections, dont on prétend qu'elles ont été influencées par un profilage avancé utilisant des données de suivi de santé en temps réel, exploitées pour manipuler les électeurs à grande échelle...

Soumise à cette connectivité omniprésente et à la dépendance aux technologies, Alice estdans l'incapacité d'échapper à l'emprise de « sa » ville intelligente. Perdue, prisonnière de son cocon numérique, elle se demande ce qui peut lui rester de sa vie privée, et même de son libre arbitre, dans cette société devenue folle.

Libérer l'optimisme et atténuer le pessimisme associés à l'IoT

L'environnement d'Alice exploite l'Internet des objets (Internet of Things, IoT), une source tant d'espoirs que de craintes, à différents niveaux. L'IoT crée de nouvelles opportunités, mais apporte également de nouveaux problèmes. Si les solutions à ces problèmes ne sont pas toutes nécessairement technologiques, science et technique peuvent certainement contribuer à contrer ce que l'IoT peut avoir de conséquences négatives. Pour libérer l'optimisme lié à l'IoT et limiter les sources de pessimisme, nous devons nous appuyer sur des progrès à réaliser dans différents domaines scientifiques, parmi lesquels :

- les **réseaux informatiques** (pour permettre aux appareils d'Alice de communiquer et d'interagir) ;
- la **miniaturisation d'équipements à faible consommation d'énergie** (pour prolonger la durée de vie des appareils d'Alice et les rendre plus pratiques) ;
- l'**intégration de logiciels économes en énergie** (pour permettre aux appareils de coopérer, durablement, à partir d'une batterie de faible capacité) ;
- l'**informatique distribuée** (pour qu'Alice puisse choisir dans une certaine mesure où et comment ses données sont traitées) ;
- un **traitement de données respectueux de la vie privée** (pour qu'Alice garde le contrôle de ses données à caractère personnel ou sensible, et de l'utilisation qui en est faite) ;
- l'ingénierie du **contrôle et la robotique** (pour piloter avec efficacité les capteurs et actionneurs qu'Alice utilise) ;
- l'**interface Homme-machine** (pour contrôler simplement le système, mais de façon performante) ;
- la **sûreté des systèmes** (pour s'assurer que les actionneurs ne sont pas dangereux pour Alice et pour les autres) ;
- la **sécurité des systèmes** (pour défendre Alice contre de possibles attaques de pirates).

Inria réunit plus de 200 équipes-projets réparties entre huit centres de recherche, et contribue à chacune de ces disciplines. Ce document présente l'apport Inria sur les principales tendances et sur les grands défis de l'IoT. Il expose l'action de ses équipes en recherche scientifique, leurs contributions au développement de logiciels et au transfert de technologies pour répondre à ces défis.

Explorant d'autres domaines, ce document identifie également les aspects par lesquels est impactée une société dépendante de l'IoT, des préoccupations éthiques aux questions de transparence, de souveraineté et d'éducation.

D'autres livres blancs ont déjà été publiés par ailleurs sur des sujets connexes. Certains mettent l'accent sur un sous-ensemble de l'IoT, comme par exemple les télécommunications et les aspects réglementation (voir les livres blancs de l'Arcep ou de l'AFNIC), d'autres sur les logiciels open source sur lesquels s'appuie l'IoT (voir le livre blanc de Systematic). D'autres encore ont examiné la question de l'IoT du point de vue d'un constructeur de matériel industriel (livre blanc de NXP) ou d'un prestataire de services informatiques (livre blanc d'Atos). Dans ce document, nous avons plutôt choisi de proposer un traitement plus global et plus fondamental de l'Internet des Objets, en nous appuyant sur les problématiques auxquelles la recherche scientifique se consacre, en la matière.

PARTIE 1

L' Internet des Objets (IoT)

Dans cette première partie, nous définissons d'abord les contours de l'IoT, sa genèse, son actualité et ses perspectives. Ensuite, nous identifions les principaux défis sociétaux, techniques et scientifiques relatifs à l'IoT.

1

IoT : Hier, aujourd'hui, demain

Si le concept est à la mode, **définir l'IoT n'est pas chose facile.** En effet, l'IoT concerne le matériel et le logiciel, les technologies réseau autant que les sciences des données, les services et le déploiement d'infrastructures, les réseaux de capteurs sans fil comme le cloud computing. L'IoT n'est-il encore qu'une vision ? Cette vision n'est-elle pas plutôt déjà une réalité ? Et pour commencer, qu'est-ce que l'IoT ? Les réponses à ces questions sont nombreuses et sujettes à débat. C'est un peu comme poser la question de la définition d'Internet.

Dans ce document, nous considérons l'IoT comme la forme tangible d'une composante importante de l'Internet de nouvelle génération. De ce point de vue, l'IoT représente un ensemble de technologies de portée générale, qui :

- jettent des ponts entre le monde numérique et le monde physique ;
- comblent l'écart entre les technologies Internet et des systèmes embarqués de plus en plus variés.

La terminologie de l'IoT n'est pas encore complètement figée. Dans ce document, nous parlons de l'IoT comme un équivalent de l'Internet du Tout (*Internet of Everything, terminologie Cisco/W3C*), de l'Internet physique (*Physical Web, Google*), de l'informatique physique (*Physical Computing, Arduino*), de la communication entre machines (*Machine-to-Machine*, M2M), des systèmes cyberphysiques (*Cyber-Physical Systems*, théorie des asservissements) ou du *World-Sized Web* (un concept proposé par Bruce Schneier).

Rapide historique de l'IoT

Avant Internet, à la fin des année soixante-dix, des produits de domotique tel X10 avaient déjà fait leur apparition sur le marché. Par la suite, le début des années quatre-vingt-dix a vu apparaître des concepts futuristes comme celui du Digital Desk, imaginant des «objets augmentés» coopérant sur le réseau au moyen d'interfaces permettant des **interactions tangibles** ouvrant des brèches dans la séparation entre les mondes numérique et physique. À peu près à la même époque, d'autres anticipaient l'Internet des Objets, dont Mark Weiser et sa vision de « l'informatique ubiquitaire et de la virtualité » incarnée, une vision qui devient peu à peu réalité.

À la fin des années quatre-vingt-dix, l'*Auto-ID Center* ouvrait l'ère des puces RFID, préfigurant un monde où quasiment chaque objet pourrait être identifié avec une adresse unique sur le réseau. Avec ce système préliminaire, chaque étiquette était dotée d'une simple puce contenant uniquement un numéro de série (l'objectif

étant d'en limiter le coût) susceptible d'être lu à proximité par une communication locale sans fil. Les données associées au numéro de série de l'étiquette étaient stockées séparément dans une base de données accessible en ligne.

Dans les années deux-mille, de nouveaux concepts et de nouvelles techniques ont permis la création de réseaux de capteurs/actionneurs sans fil (WSAN ou WSN) : de minuscules ordinateurs alimentés par batterie, collaborent initialement pour établir des réseaux sans fil (à sauts multiples) avant d'utiliser ces réseaux pour transporter les données de leurs capteurs, ou pour transmettre les commandes des actionneurs.

Donnant à ces notions une portée plus générale, le terme « informatique omniprésente » (*pervasive computing*, terme qui se rapproche du concept d'informatique ubiquitaire) capture la tendance consistant à embarquer dans les objets du quotidien des capacités de calcul et de communication. Plus ou moins synonyme, l'utilisation du terme Internet des objets (*Internet of Things, IoT*) a commencé à se généraliser dans les années deux-mille dix.

Ces dix dernières années ont vu apparaître de plus en plus d'objets augmentés, intégrant différents niveaux de puissance de calcul et de coopération sur le réseau..

Les innovations permettant l'émergence de l'IoT

L'émergence de l'IoT a connu récemment une accélération grâce aux innovations réalisées dans les domaines des matériels embarqués et des réseaux basse consommation, des logiciels systèmes embarqués et de l'informatique en périphérie de réseau (*edge computing*). Des industriels issus des horizons les plus variés (de la PME aux *Big Tech*) participent à l'innovation dans ces domaines, à différents niveaux. De nouveaux organismes de normalisation ont vu le jour, ainsi que de nouveaux standards technologiques pour l'IoT. Les lecteurs intéressés trouveront ci-dessous des références concrètes, dont la liste est loin d'être exhaustive : notre objectif n'est pas d'établir la pertinence de tel ou tel acteur, mais plutôt d'illustrer la diversité de ces innovations.

Réseau de capteurs FIT/IoT Lab installés dans le centre Inria Grenoble Rhône-Alpes.
© Inria / Photo H. Raguet.

Innovation dans le matériel embarqué

Les ordinateurs monocartes à bas coût sont aujourd'hui courants (citons par exemple le *RaspberryPi* ou le *Jetson Nano d'NVIDIA*...) basés sur des microprocesseurs de plus en plus performants. D'autre part, des entreprises comme *STMicroelectronics, Microchip Technology Inc., Espressif* ou *SiFive* développent de nouveaux appareils IoT à très basse consommation d'énergie, en s'appuyant sur de nouvelles architectures de microcontrôleurs conçues par des sociétés telles que *ARM Ltd.* ou en intégrant des normes matérielles *open source* du type RISC-V. À cela s'ajoutent de nouveaux coprocesseurs de sécurité et de cryptographie économes en énergie, disponibles par exemple auprès de sociétés comme NXP *Semiconductors ou Nordic Semiconductors.*

Innovation dans les réseaux basse consommation

De petits appareils, de la taille d'un capteur/actionneur, sont mis en réseau à l'aide de nouvelles radios basse consommation (protocoles LoRa, 802.15.4, BLE...), et de la miniaturisation des piles de protocoles permettant un réseau polyvalent (6LoWPAN) s'appuyant sur un pool d'adresses réseau uniques quasiment infini (avec IPv6). On assiste aussi à l'émergence de matériels de communication sans batterie, à récupération d'énergie, sous l'impulsion d'entreprises telles que *EnOcean* ou *Onio*. Des consortiums et des organismes de normalisation comme *LoRa*

Alliance, IEEE, IETF, W3C ou *3GPP* produisent de nouvelles spécifications ouvertes pour les protocoles de communication réseau prenant en charge les dispositifs à faible consommation d'énergie. De grandes entreprises comme *Semtech* et *Texas Instruments* fournissent de nouvelles puces radio économes en énergie. Ces puces sont utilisées par les vendeurs de matériel et par de nouveaux opérateurs comme *Sigfox* ou *Actility*, qui se concentrent sur les technologies IoT (et par les grands opérateurs classiques tels qu'Orange, SFR ou Bouygues).

Innovation dans les logiciels embarqués

Les distributions compactes embarquées de Linux se sont imposées comme étant les plates-formes logicielles de choix pour les ordinateurs monocartes (des appareils IoT construits autour de microprocesseurs « classiques »), qui bénéficient des contributions *open source* d'un grand nombre de développeurs d'entreprises de tailles variables. Sur les plus petits équipements (des appareils IoT bas de gamme, basés sur des microcontrôleurs), de nouvelles plates-formes logicielles de systèmes embarqués dont le code est *open source* tirent profit du développement collaboratif de logiciels basse consommation, comme *FreeRTOS (Amazon), Zephyr (Intel), Arduino* ou *RIOT*.

Innovation dans l'informatique en périphérie de réseau

Cette discipline vise à rapprocher physiquement la source des données et les infrastructures de calcul et de stockage de données. De nouveaux écosystèmes accélèrent le déploiement, l'exploitation et la maintenance de l'informatique en périphérie de réseau. Par exemple, le développement de logiciels d'apprentissage automatique embarqués est facilité par des plates-formes telles que *TFLite*, une plate-forme pilotée par *Google*. Les nouveaux coprocesseurs matériels et logiciels de sociétés comme *ARM* ou *Greenwaves Technologies* améliorent la capacité opérationnelle des réseaux neuronaux sur les dispositifs IoT situés à la périphérie du réseau. L'informatique en périphérie de réseau bénéficie également de l'émergence de nouveaux outils de gestion *open source* dédiés à l'IoT comme ceux que développe l'organisation *Eclipse IoT Foundation*, ou des outils de gestion de flotte comme le système *Kubernetes*. Les services de *cloud* prennent en charge l'IoT et l'informatique en périphérie avec des services adaptés, notamment ceux fournis par *Microsoft* avec *Azure IoT Hub* ou *Amazon* avec *AWS IoT GreenGrass*.

Le déploiement de l'IoT est en plein essor, tandis que l'intégration de ces différentes innovations est facilitée par les technologies internet génériques et les infrastructures partagées dans le *Cloud*.

Au-delà de la recherche, Inria a également engagé une démarche systématisant le soutien à des projets *Deep Tech* qui innovent dans le domaine de l'IoT en appliquant les résultats de la recherche. Le programme Inria Startup Studio accompagne ainsi la création de projets de startups innovantes. De jeunes pousses telles *Falco, Stackeo, Statinf* ou *CryptoNext* sont des exemples récents de ces créations d'entreprises dans le domaine de l'IoT.

FALCO est une entreprise issue d'Inria qui propose du matériel IoT, ainsi que du logiciel et des services dédiés à la gestion environnementale des ports maritimes de plaisance : suivi en temps réel des emplacements disponibles, contrôle optimisé des ressources, lutte contre la pollution, sensibilisation aux bonnes pratiques. Les produits de Falco s'appuient notamment sur la technologie IoT sans fil à basse consommation développée par les chercheurs d'Inria.

STACKEO est un éditeur de logiciels issu d'Inria, qui aide les entreprises à industrialiser à l'échelle souhaitée leurs solutions IoT et leur connectivité. Construite sur un business model de type logiciel en tant que service (SaaS), Stackeo propose une suite d'outils pour mettre en œuvre une stratégie IoT, pour faire travailler ensemble les équipes des systèmes informatiques et des technologies opérationnelles (IT/OT), et pour piloter des chaînes de valeur durables. Stackeo développe le concept d'architecture IoT en tant que code, à partir d'un langage de modélisation dédié et d'une méthodologie systémique brevetée.

STATINF est une société fournissant des outils permettant l'analyse statistique du comportement temporel des logiciels embarqués, pour les systèmes temps réel impliquant des processeurs multicoeurs. Avec de tels outils, les systèmes embarqués hautement critiques des industries telles que l'avionique, l'automobile, les drones, l'aérospatiale peuvent interpréter les variations d'exécution possibles du logiciel qui utilise leur matériel IoT embarqué.

CRYPTONEXT est une entreprise dans le domaine de la cybersecurité, issue d'Inria et de l'Université de la Sorbonne. Cette entreprise fournit des bibliothèques logicielles mettant en œuvre des algorithmes cryptographiques conçus et optimisés pour la sécurité postquantique. CryptoNext fournit également fournit également des conseils dans le domaine de la cybersécurité résistante aux attaques quantiques.

L'IoT aujourd'hui : le point de bascule

Partons du principe que les appareils IoT sont les appareils connectés de tout type, grand public et B2B, à l'exclusion des téléphones, tablettes, ordinateurs portables et de bureau. Dans ce cas, entre 2010 et 2020, le nombre d'appareils IoT connectés a connu une augmentation sidérante, de l'ordre de 1 000 %. Près de 10 milliards d'appareils IoT ont été mis en service et interconnectés ces dix dernières années. Tentons de mettre en perspective l'ampleur de cette évolution : en 2010, les appareils IoT représentaient 10 % du total des appareils connectés. En 2020, cette part atteignait plus de 50 % ; en d'autres termes, nous avons dépassé un point de bascule. Désormais, les appareils IoT sont officiellement plus nombreux que les appareils non-IoT. Selon un rapport d'analyse du marché mondial, l'IoT représentait en 2019 un chiffre d'affaires global de plus de 300 milliards de dollars, et ce chiffre augmente rapidement.

Beaucoup d'appareils IoT intègrent des microcontrôleurs basse consommation. Plus de 28 milliards de microcontrôleurs ont été expédiés en 2018, et on estime qu'en 2020, plus de 250 milliards de microcontrôleurs étaient utilisés dans le monde. Si tous les microcontrôleurs ne sont pas connectés en réseau, de plus en plus de déploiements combinent des microcontrôleurs exécutant un code de plus en plus complexe et une connexion directe ou indirecte à un réseau. La majorité d'entre eux se connecte au moyen d'un réseau sans fil (personnel, local, étendu ou réseau cellulaire). Ces déploiements se multiplient sur les segments les plus variés. Notamment (par ordre décroissant de parts de marché), les applications automobiles, l'automatisation industrielle, les appareils grand public personnels/domestiques, les compteurs intelligents intégrés aux *smart grids*, les applications de santé, l'aérospatiale et la défense. Dans le secteur industriel, par exemple, le réseau concerne à présent une multitude de processus qui reposaient jusqu'alors sur une logique de contrôle au moyen de capteurs et d'actionneurs locaux. Les données agrégées en temps réel et les « jumeaux numériques » permettent de mieux apprécier dans sa globalité la complexité des chaînes d'approvisionnement, de production et de distribution, et de mieux les contrôler à l'échelle mondiale

Quelles perspectives pour l'IoT ? Des milliards de nouvelles applications pour détecter et agir

Des déploiements IoT sont envisagés sur la quasi-totalité des marchés et secteurs d'activité : maison intelligente, bâtiments à énergie nette zéro, e-santé, industrie 4.0, agriculture de précision, surveillance en temps réel de la faune et de l'environnement, villes intelligentes, chaîne logistique du transport de marchandises, partage de voitures, de vélos et de scooters... Plutôt que de tenter de dresser la liste exhaustive de ces applications, nous renvoyons le lecteur intéressé à des lectures complémentaires. Les analystes prédisent que des dizaines de milliards d'appareils IoT supplémentaires, connectés au réseau, seront déployés partout sur la planète et qu'ils éclipseront rapidement toute connexion non IoT.

D'un certaine manière, l'IoT donne « bras et jambes » à Internet pour évoluer dans le monde physique. L'IoT fait naître de nouvelles possibilités de communication, de détection (collecte de données), de raisonnement (utilisation de ces données), puis d'action physique au moyen d'actionneurs. Des boucles de contrôle d'un type nouveau, reposant sur le réseau, peuvent exploiter si nécessaire d'importantes ressources de calcul distantes. En rendant possibles ces nouvelles boucles de contrôle, puis en les mettant en œuvre, l'IoT transforme radicalement l'environnement :

Les **capteurs** permettent une plus fine observation de nombreux processus. Ils peuvent par exemple détecter un dysfonctionnement au sein d'un système complexe, et contribuent à réduire considérablement les dépenses d'exploitation. Les leaders industriels de l'IoT projettent d'optimiser les processus industriels en combinant le matériel, la connectivité sans fil IoT et des flux de données complexes : les problèmes et les pannes peuvent être détectés beaucoup plus rapidement (en quelques minutes ou quelques heures, quand cela demande aujourd'hui encore des jours, voire des semaines). Cette efficacité fait économiser des centaines de milliers de dollars gaspillés par an et par usine. L'augmentation du rendement constant autorise alors une croissance à deux chiffres de la production et du bénéfice brut.

Les **actionneurs** permettent une (re)configuration rapide, adaptative et automatisée de systèmes cyberphysiques complexes. On peut par exemple déployer des stratégies autonomes d'économie d'énergie à différents niveaux (un ensemble de bâtiments ou une ville entière), qui interagissent de manière dynamique pour obtenir des réductions spectaculaires de la consommation d'énergie. On s'attend

également à ce que l'IoT réduise notre impact sur l'environnement (par la détection de la pollution et une meilleure prise en compte des consommations d'énergie), et que cet effet vertueux l'emporte globalement sur le coût environnemental de la production, du déploiement et de la maintenance de l'IoT.

Avec les **boucles de contrôle cyberphysiques** basées sur le traitement des données des appareils IoT à différentes échelles de temps, il est possible d'envisager de nouveaux niveaux de contrôle et de prévention. Par exemple, en associant télémesure de pointe et apprentissage automatique, on peut effectuer une maintenance prédictive bien avant l'apparition réelle de problèmes et de pannes sur un automate industriel. Plus que la simple amélioration des processus existants, les boucles de contrôle cyberphysiques devraient créer des processus et des services totalement nouveaux, dont l'impact sur la société sera important

Les **composants robotiques** sont à la fois une conséquence et un élément contributeur des déploiements de l'IoT. D'une part, l'association de capteurs et d'actionneurs, puis leur mise en relation avec des boucles de contrôle cyber-physiques, créent des systèmes de type robot. D'autre part, des flottes de robots, de drones et d'autres dispositifs autonomes sont déployées pour compléter l'infrastructure IoT si nécessaire ; ces dispositifs peuvent recueillir des données et contribuer à la fourniture dynamique de ressources.

2

Les défis sociétaux de l'IoT

L'IoT a déjà commencé à transformer notre société, en y introduisant un certain nombre de changements fondamentaux. Dans ce chapitre, nous proposons une vue d'ensemble des principaux questionnements soulevés dans ce domaine. Nous offrons également un aperçu de solutions qu'Inria apporte au débat.

Le cadre juridique : l'équilibre entre innovation sans permis et soucis éthiques

Héritier d'Internet, l'IoT connaît une croissance extrêmement rapide et fait l'objet d'intérêts souvent contradictoires.

D'autre part, l'utilisation de ces technologies suscite des inquiétudes de plus en plus marquées sur le plan éthique, par exemple en ce qui concerne la vie privée ou l'environnement.

L'utilisateur moyen de l'IoT ne saisira vraisemblablement pas dans leur complexité les implications de son utilisation de ces appareils connectés. Il est donc essentiel de concevoir des cadres juridiques appropriés pour guider le développement de l'IoT.

Il reste que ces innovations sont généralement adoptées plus rapidement que ne sont mises en place la législation et la réglementation pour les encadrer, ce qui crée des zones grises et des failles réglementaires. Compte tenu du « taux de pénétration » attendu de l'IoT, ces failles pourraient se révéler vraiment problématiques si l'on considère l'ampleur, la nature et le niveau de détail des données collectées.

Il faut donc agir avec prudence pour limiter cet impact négatif autant que possible sans entraver l'innovation. En la matière, ce sont principalement les gouvernements et les organismes de réglementation qui ont la capacité et la responsabilité d'agir : il suffit d'observer l'impact du RGPD pour se convaincre de l'intérêt d'une action concertée au niveau européen.

La confiance de l'opinion publique : la gagner et la conserver

Pour accélérer l'adoption de l'IoT, il faut davantage de transparence et donner à l'utilisateur final les moyens de se faire entendre des gouvernements et des industriels.

Les déploiements IoT qui réunissent plusieurs parties prenantes sont souvent motivés par des intérêts qui finissent par diverger avec le temps. En l'absence d'instruments techniques ou juridiques adaptés, comment gérer ces divergences de façon juste et appropriée ? Plus particulièrement, les consommateurs sont plus exposés dans le rapport aux professionnels à une relation B2C (plutôt qu'à une relation B2B) et doivent donc être protégés davantage.

Captation via téléphones portables du bruit ambiant pour établir une carte collaborative de la pollution sonore. © Inria / Photo C. Morel.

Le public est de plus en plus conscient de la façon dont l'État utilise les technologies Internet pour organiser la surveillance de masse (voir le scandale Snowden). Il connaît désormais les techniques du piratage en ligne motivé par le profit, ou l'exploitation des données par le capitalisme de surveillance et la personnalisation avancée des publicités.

Des critiques, voire des réactions de rejet, prennent forme dans la sphère publique à l'encontre de projets pilotes de déploiement plus ambitieux de l'IoT, qui pourraient favoriser une certaine opacité ou des monopoles industriels de fait. Dans le même temps, la confiance des utilisateurs s'érode à mesure que les controverses s'embrasent, dès qu'est révélée une nouvelle faille de sécurité (souvent très élémentaires !) touchant d'innombrables appareils IoT, ou que l'on découvre de nouveaux processus menaçant la vie privée et des fonctionnalités cachées des dispositifs. Il faut s'attendre à de nouvelles flambées de controverses, par exemple au sujet des compromis envisagés entre vie privée et nécessité de surveillance, pour les gouvernements, dans un contexte de protection de la sûreté nationale.

L'une des principales caractéristiques d'Internet, et probablement l'un de ses meilleurs atouts à ce jour, a été d'accorder de l'autonomie à l'utilisateur final, de lui donner les moyens de faire des choix qui ne soient pas dictés par un fournisseur ou par un gouvernement. À la lumière des récentes tendances politiques et techniques, on peut toutefois craindre que la responsabilisation de l'utilisateur final et la transparence ne perdent du terrain. Dans ce contexte, l'IoT doit impérativement contribuer à restaurer et réaffirmer ces principes, et par conséquent, à regagner la confiance du public. Si l'IoT n'est pas en capacité de le faire, son déploiement pourrait ne pas suivre les prévisions

Souveraineté

À défaut d'être maîtrisé, l'IoT pourrait accroître la dépendance envers une ou plusieurs technologies, qui ne seraient ni neutres sur le plan géopolitique, ni compatibles avec le respect de la vie privée.

La technologie occupe une place de plus en plus centrale dans la vie des gens. Les questions de souveraineté deviennent donc essentielles, tant à l'échelle individuelle qu'au niveau de l'État.

- **Pour l'individu,** la difficulté consiste à maximiser et conserver sa capacité à s'affranchir de la dépendance envers des fournisseurs et des technologies susceptibles de menacer la vie privée. Par exemple, afin d'interagir avec ses appareils IoT, un utilisateur averti pourrait préférer utiliser une machine intermédiaire qu'il contrôle totalement pour réaliser un paramétrage, plutôt qu'une configuration dont l'intégration est plus «fluide» mais passe par un service cloud. Les grandes entreprises qui entretiennent chacune leur « jardin exclusif » pourraient décourager ce type d'approches, voyant dans ce cas le client comme un assaillant qu'il faut contrôler.

- **Pour les États,** le défi consiste à s'efforcer d'être le moins dépendant possible de solutions techniques susceptibles d'être utilisées de manière hostile. Prenons un exemple concret : dépendre de *Linux* peut être considéré comme plutôt neutre d'un point de vue géopolitique, alors que la décision de s'appuyer sur Android est nettement moins neutre (même si techniquement, *Android* est basé sur *Linux*). Il n'est pas toujours facile de prendre de telles décisions, surtout dans des domaines que l'on a longtemps négligés (ou sous-traités).

Quoi qu'il en soit, rien n'est gratuit en ce monde. Gagner en souveraineté implique une forme de sacrifice. Pour les individus, cela se concrétise typiquement par une dégradation de l'aspect pratique de ces usages, au moins temporairement. Pour les États, la contrepartie est comparable à ce que l'on retrouve dans le monde des affaires : l'indépendance exige généralement des investissements nettement plus conséquents en recherche, développement et maintenance. En outre, si cette souveraineté doit être gagnée au prix du développement et de l'utilisation d'un plus grand nombre de systèmes distincts et concurrents, d'autres difficultés, concernant l'interopérabilité, par exemple, ne manqueront pas de se présenter. Le tout est donc d'optimiser l'investissement, pour « en avoir pour son argent » autant que possible.

Normalisation

À mesure que la technologie IoT se fait toujours plus présente dans la société et dans notre quotidien, il devient de plus en plus urgent de définir des normes pour l'encadrer.

Par essence, la technologie IoT consiste à permettre à des systèmes embarqués, hétérogènes, de se connecter et d'interagir sur le réseau, potentiellement à grande échelle. Sans surprise, la normalisation est donc un aspect central de l'espace IoT, et plus particulièrement la normalisation des communications du réseau des appareils IoT. Par exemple, les organismes de normalisation comme le *World Wide Web Consortium* (W3C), l'*Internet Engineering Task Force* (IETF) ou l'*Institute of Electrical and Electronics Engineers* (IEEE) sont autant de forums incontournables où sont prises des décisions techniques majeures, dont les ramifications sont infinies, y compris au sein de la société.

Dans la pratique, chaque organisme de normalisation porte des valeurs qui lui sont propres, et cette absence de neutralité se reflète donc dans la normalisation de l'IoT. En particulier, le pouvoir concentré des *Big Tech* sur la société, combiné à leur influence décisive sur le développement des prochaines normes (par un processus de normalisation de facto) crée des situations auxquelles une réponse

doit être apportée. Il faut du reste mettre en lumière les rapports qui existent entre les normes de la technologie IoT et les droits humains, par exemple. La difficulté consiste ainsi à faire mieux connaître à l'opinion publique le fonctionnement des technologies standard, mais aussi l'impact que différentes caractéristiques de ces technologies peuvent avoir sur la société.

Éducation

Plus l'IoT met en réseau des ordinateurs «invisibles», des capteurs et des actionneurs utilisés inconsciemment, plus il devient nécessaire d'inclure dans les programmes d'enseignement les principes de base des systèmes cyberphysiques, et de présenter ce qu'ils peuvent avoir de dangereux.

Si l'enseignement de l'informatique en général est complexe, c'est d'autant plus le cas pour l'IoT. Ceci s'explique en grande partie parce que la technologie IoT est encore extrêmement fragmentée, en particulier pour l'IoT bas de gamme. Ce défi comporte plusieurs aspects :

- ***Développer les compétences techniques requises :*** sur un plan purement technique, les difficultés de l'enseignement sont évidemment exacerbées dans les domaines où l'enseignement technique générique de l'informatique a pris du retard. Certaines compétences techniques requises pour l'IoT, comme la programmation de logiciel embarqué, sont encore trop rares.
- ***Développer les compétences comportementales nécessaires :*** sur un plan moins technique, il est toujours souhaitable de posséder un certain niveau d'éducation et de « sensibilisation » pour développer le bon sens et l'esprit critique concernant ce que les produits IoT peuvent/doivent (ou ne peuvent/ne doivent pas) faire. Cette capacité de réflexion critique est peut-être en passe de devenir rare, et cela peut constituer un changement majeur pour notre société.

Notre appréhension de la réalité, dans un monde où l'IoT est partout, est en soi un sujet d'étude pour la philosophie. Les capteurs et actionneurs nous permettent d'individualiser et de vivre différemment non seulement la réalité virtuelle, mais aussi la réalité physique. D'un côté, des disciplines comme la postphénoménologie s'efforcent de caractériser l'interaction entre les humains, le monde naturel et les technologies modernes, alors que ces dernières tiennent un rôle de médiateur qui se révèle de moins en moins neutre. D'un autre côté, on travaille à la conception de systèmes « sociotechniques », prenant en compte non seulement le matériel numérique et technologique, mais aussi les utilisateurs humains eux-mêmes, considérés comme éléments du système et « matériau ».

Lutter contre le changement climatique et la raréfaction des ressources

Si l'IoT peut être un outil de la lutte contre le changement climatique, la multiplication de gadgets IoT peut en revanche contribuer à l'épuisement des ressources.

À l'échelle mondiale, le déploiement de milliards d'appareils IoT consommera d'énormes quantités d'énergie et de ressources (plastiques, métaux, batteries…), pour la production tout d'abord, puis pour l'acheminement et l'exploitation de ces appareils. Dans notre contexte de crise écologique, on doit s'interroger sur la pertinence de ces équipements.

Prévoir les événements de gel dans les vergers de pêchers. © Inria / Photo G. Scagnelli.

En principe, l'IoT fournit des composants et des outils importants, dont on a besoin pour suivre avec précision et en temps réel le changement climatique et l'incidence des politiques mises en place pour l'enrayer. Plus encore, l'IoT peut contribuer à l'automatisation des ajustements dynamiques nécessaires, et à mettre en œuvre des optimisations de nombreux systèmes complexes et grands consommateurs de ressources et d'énergie, comme les *workflows* industriels ou les systèmes de gestion de l'énergie des maisons et bâtiments intelligents. L'IoT peut indéniablement être utile pour lutter contre le changement climatique. Cependant, nous devons continuer d'étudier son efficience nette sur le plan de l'énergie et des ressources consommées, et mettre en place des contraintes juridiques évolutives pour orienter cette efficience.

Comment Inria contribue à répondre aux défis sociétaux de l'IoT

Inria est un établissement public scientifique et technologique (EPST) qui fournit son expertise scientifique dans différents domaines liés à l'IoT et qui produit de nombreux travaux de recherche appliquée, exerçant une influence sur l'éducation, la souveraineté, les normes et les cadres juridiques. Inria contribue ainsi à gagner la confiance du public concernant l'utilisation des nouvelles technologies IoT.

Dans la pratique, Inria contribue par différents aspects au maintien de la **souveraineté** en lien avec l'IoT. L'institut mène des travaux de recherche scientifique que ses pairs considèrent comme excellents, et en publie systématiquement les résultats sur des sites en libre accès, contribuant ainsi à rendre la souveraineté possible. La contribution d'Inria est toutefois plus large que cet apport scientifique.

En plus de ses travaux de recherche, Inria participe à de nombreux projets internationaux, des collaborations techniques et transdisciplinaires. *Via* ses équipes de chercheurs, Inria supervise le déploiement de nouvelles technologies IoT dans le cadre de différentes applications concrètes.

D'autre part, Inria produit, édite et entretient des logiciels *open source*. D'importantes communautés d'utilisateurs se sont développées autour de projets de logiciels open source à fort impact pilotés par Inria (*RIOT, scikit-learn*, etc.). Ces « briques logicielles » constituent des plates-formes, sur lesquelles un système peut s'appuyer pour être performant sans compromettre sa neutralité géopolitique.

La manière dont Inria prête son expertise scientifique aux actions de **normalisation** est un autre exemple de l'apport de l'institut en matière de **souveraineté** liée à l'IoT. Inria participe ainsi régulièrement à l'élaboration de nouvelles spécifications de **normes ouvertes** pour la technologie IoT, publiées par de grands organismes de normalisation comme *IETF* et *W3C*.

Inria utilise également son expertise dans divers domaines scientifiques liés à l'IoT, contribuant au débat sur **le cadre juridique et les garde-fous** guidant l'IoT. Par exemple, les recherches entreprises chez Inria portent notamment sur l'évaluation des problèmes pratiques de mise en conformité avec la réglementation européenne RGPD, posés par les produits IoT grand public dans le secteur de la maison intelligente.

Inria contribue activement à l'**éducation** dans le domaine de la technologie IoT, à différents niveaux. De nombreux chercheurs à l'institut sont également enseignants, notamment à l'université. Mais Inria contribue également à l'**éducation**

au-delà de l'implication traditionnelle des enseignants-chercheurs, en proposant par exemple des formations en ligne ouvertes à tous (*Massive Open Online Course*, MOOC) adaptées à un large public dans le domaine de l'IoT. Ces formations sont suivies par des milliers de participants en Europe et à travers le monde. Inria a aussi conçu des programmes éducatifs destinés aux plus jeunes, qui donnent à des jeunes de quinze ans les connaissances dont ils ont besoin pour mieux appréhender la technologie numérique. Enfin, l'implication d'Inria dans les activités pédagogiques prend aussi la forme de la publication d'une série de livres blancs, dont le présent document fait partie.

3

Les défis scientifiques et techniques de l'IoT

Nous avons les attentes les plus diverses à propos de l'IoT : autonomiser les individus grâce à des services révolutionnaires, exercer une influence majeure sur la société et l'industrie, voire sauver la planète en luttant contre le changement climatique à l'échelle mondiale.

Quels défis scientifiques devrons-nous relever pour répondre à de telles attentes ? Voici une liste de questions, dont nous considérons qu'elles constituent des enjeux essentiels pour l'IoT.

Comment concilier sphère privée et IoT omniprésent ?

Il existe en permanence des tensions entre la volonté d'exploiter les données IoT et la nécessité de respecter la vie privée des utilisateurs de l'IoT. À un extrême, il **devrait** être possible d'exploiter massivement les données des utilisateurs d'appareils IoT s'il s'agit de sauver des vies. À l'autre extrême, le déploiement tous azimuts de l'IoT **pourrait** fournir les éléments clés d'un « panoptique » numérique. Un tel dispositif pourrait réduire à néant la sphère privée des individus. Se pose donc une question cruciale : comment, et dans quelle mesure, pouvons-nous garantir une protection forte de la vie privée tout en préservant l'utilité des données IoT ?

La difficulté est multiple. Elle concerne d'une part, la conception de nouveaux paradigmes et de nouvelles techniques de prétraitement des données, applicables directement sur les appareils IoT, obfusquant les caractéristiques des données qui ne répondent pas à un « intérêt légitime » précis. D'autre part, le problème est également de concevoir de nouvelles primitives cryptographiques capables de résister à des attaques postquantiques, tout en étant assez peu gourmandes en ressources pour pouvoir s'exécuter sur les appareils IoT.

Comment renforcer la résilience, la sûreté et la sécurité de l'IoT ?

À mesure que nous devenons dépendants de nouveaux services créés à partir de l'IoT, sa résilience face à une panne partielle de l'infrastructure ou à un dysfonctionnement des sous-systèmes devient un élément crucial. Les déploiements de l'IoT impliquent des architectures de systèmes distribués de plus en plus complexes, pour lesquelles obtenir une résilience « par conception » est une difficulté majeure. Le degré de fiabilité joue également un rôle en ce qui concerne la sûreté et la sécurité de l'IoT.

On considérait jusqu'alors que la cybersécurité concernait principalement le cyberespace : sécuriser ses informations numériques, respecter la « netiquette » (les bonnes pratiques de l'usage d'Internet), etc. Avec l'IoT, la cybersécurité déborde de l'espace virtuel pour gagner l'espace physique : il s'agit désormais de garantir son intégrité physique et de protéger son environnement dans le monde réel.

Map of Things : collecte des données et information des usagers d'objets connectés. © Inria / Photo C. Morel.

À quels risques nouveaux s'expose-t-on en matière de sécurité ou de sûreté avec l'IoT ? Pour quels avantages ? De simples « gadgets » IoT ne valent peut-être pas les risques qui sont indissociables d'eux. Il faut développer de nouveaux modèles pour comprendre qui sont les attaquants potentiels et quels sont les enjeux de sécurité, dans un environnement IoT complexe. Ensuite, on pourra développer à partir de ces modèles de nouveaux mécanismes de garantie sur les logiciels, le matériel et les communications des appareils IoT, tout au long d'une durée de vie qui peut s'étendre à plusieurs dizaines d'années. Ce défi est d'autant plus saillant pour les dispositifs IoT basse consommation, qui sont le nouveau « maillon faible » de cet environnement.

Comment associer apprentissage automatique et IoT ?

On développe des applications associant intelligence artificielle et apprentissage automatique (*Machine Learning*) dans les domaines les plus variés. D'un côté, l'IoT fournit en grandes quantités les précieuses données nécessaires à l'entraînement

des modèles d'apprentissage automatique. De l'autre, l'IoT utilise des capacités d'inférence, laquelle repose sur l'utilisation de ces modèles préentraînés.

L'important est de savoir comment exploiter davantage de données IoT tout en réduisant l'impact sur la vie privée et la charge sur le réseau. L'une des difficultés sur ce point est de concevoir des alternatives au modèle centralisé pour l'apprentissage automatique : des techniques d'entraînement distribué, de manière fédérée, entre pairs. Dans quelle mesure l'apprentissage distribué peut-il être performant et fiable ? Jusqu'où peut-on abaisser le niveau des ressources exigées dans cet apprentissage ? À l'inverse, la difficulté est aussi de savoir comment adapter les capacités d'inférence à des appareils IoT plus petits tout en limitant les pertes de performance, et en conservant une souplesse permettant d'adapter progressivement les modèles déployés sur ces appareils, tout au long de leur durée de vie.

Comment étendre la connectivité de la boucle locale pour l'IoT ?

L'IoT dépend avant tout de la connectivité des milliards de nouveaux appareils situés en périphérie du réseau. Pour faire face à la très forte augmentation du trafic qui découle de ces connexions, on a besoin de protocoles réseau de nouvelle génération, capables de gérer la rareté des moyens de communication et d'améliorer les capacités entre appareil et infrastructure et entre appareils.

Avec un si grand nombre d'appareils, dont certains de très faible puissance, les technologies réseau de pointe utilisées pour la boucle locale (last hop) doivent être extrêmement abordables, à l'étape de l'investissement autant qu'en phase d'exploitation. Comment promouvoir une dynamique comparable à celle d'Internet, où *« la connectivité en soi est la récompense » ?* Dans le même temps, on attend une meilleure pénétration en intérieur, une diminution des besoins en énergie, ainsi que des communications à plus longue portée en extérieur pour toucher des endroits plus reculés. Les termes « boucle locale » et « appareil IoT » devront également couvrir le domaine spatial, comme le souligne l'intensification des activités commerciales à laquelle on assiste actuellement dans l'espace orbital et au-delà.

Comment repousser les limites des concepts réseaux de bout en bout pour l'IoT ?

Internet s'est développé rapidement sur le principe voulant que l'intelligence soit placée aux extrémités plutôt que cachée à l'intérieur du réseau. L'IoT vient remettre en question ce principe. Dans quelle mesure peut-on élargir les fondamentaux du réseau de bout en bout ?

Les années 2010 ont vu apparaître des mécanismes reprenant les principes de la mise en réseau de bout en bout, comme les protocoles réseau IPv6 selon la norme ouverte 6LoWPAN, ou les travaux préliminaires sur le Web des objets (*Web of things, WoT*), qui donnent un aperçu de ce que pourrait être demain l'IoT de bout en bout. Il reste toutefois des obstacles à la finalisation d'une architecture généraliste adaptée à l'IoT. Déléguer certaines parties de l'intelligence à des intermédiaires semble inévitable pour les appareils IoT de très faible puissance. Mais comment faire ? Et dans quelle mesure faut-il déléguer ?.

On aura du reste besoin de nouveaux protocoles et d'une nouvelle sémantique pour automatiser des niveaux supplémentaires de communication et d'interopérabilité de machine à machine.

De quelles interfaces Homme-machine l'IoT a-t-il besoin ? Et quelles interfaces découlent de l'IoT ?

D'un côté, l'IoT ajoute des milliards de nouvelles interfaces vers le monde physique, depuis le cyberespace. À l'inverse, comment les utilisateurs humains peuvent-ils se servir de ces interfaces IoT pour mieux interagir avec le cyberespace ?

Il est essentiel de pouvoir améliorer tant l'accessibilité que le contrôle *via* ces interfaces. L'une des difficultés dans cette démarche consiste à identifier le niveau de contrôle adéquat pour fournir aux humains une autonomie dans leur utilisation. Les humains constituent une foule hétérogène, et n'ont pas tous les mêmes capacités pour comprendre les mécanismes à l'œuvre dans l'IoT, ni la même exigence sur ce point. Un autre aspect consiste à concevoir une nouvelle ergonomie capable de matérialiser physiquement certains éléments du cyberespace. La difficulté est d'aider les utilisateurs à apprendre, à comprendre et à s'approprier plus simplement la technologie de l'IoT, tout en accompagnant son développement et son évolution dans la durée.

Comment créer des passerelles entre IoT, contrôle et robotique ?

De nouveaux types de boucles de contrôle apparaissent à mesure que l'IoT connecte des capteurs et actionneurs à des capacités de calcul, accessibles sur le réseau, localement ou à distance. Quelles sont les boucles exploitables ? Comment paramétrer le contrôle ?

D'un côté, la recherche dans le domaine de l'IoT industriel s'efforce d'établir et d'optimiser le contrôle exercé sur des chaînes d'approvisionnement et de production d'une extrême complexité. La modélisation et le contrôle de systèmes hybrides distribués associant état continu et événements discrets est une difficulté en soi.

D'un autre côté, on peut voir les microrobots et la robotique en essaim comme une nouvelle frontière, qui impose de résoudre simultanément de nombreuses questions ouvertes en recherche : la prédictibilité de la latence, la mobilité, la localisation et l'économie de ressources conforme à l'exigence de basse consommation.

Comment concevoir des appareils IoT miniaturisés à l'échelle du millimètre?

Les évolutions récentes de la microélectronique repoussent les limites de la miniaturisation : des prototypes de puces de la taille d'un grain de riz (voire moins) sont capables de détecter, de calculer et de communiquer par liaison sans fil, sans avoir besoin de composants complémentaires. Cette « poussière intelligente » (*"smart dust"*) a le potentiel de révolutionner les microappareils portables et la robotique en essaim, mais s'accompagne de challenges spécifiques, que ce soit pour l'interopérabilité avec les standards de communication sans fil à faible puissance, la programmation du logiciel embarqué ou l'étalonnage.

Comment tendre vers la neutralité pour l'empreinte de l'IoT sur les ressources naturelles ?

L'IoT peut être un outil pour mettre en œuvre des politiques environnementales, ou pour mesurer leurs effets. Mais qu'en est-il de l'empreinte de l'IoT sur les ressources ?

Pour réduire l'empreinte de l'IoT et rendre acceptable le coût du déploiement massif apparaissant nécessaire, il est essentiel d'abaisser autant que possible le coût environnemental de la production et de l'exploitation des différents appareils. Une approche est de produire les composants électroniques en intégrant au processus nettement moins de ressources non renouvelables comme le plastique et le métal. Une autre est de concevoir de nouveaux matériels embarqués, ainsi que des logiciels et des paradigmes de mise en réseau capables d'exploiter une énergie ambiante intermittente, et d'ainsi proposer de nouveaux compromis entre performance et énergie consommée.

Par ailleurs, évaluer l'impact net global de l'Internet des objets reste difficile. Dans l'ensemble, quels sont les bénéfices nets de l'IoT et pour quelle empreinte ? Il est nécessaire de déployer un travail interdisciplinaire complexe pour établir un tableau complet des impacts directs et indirects, de l'ensemble des cycles de vie et des effets induits.

Comment Inria contribue à répondre aux défis scientifiques et techniques de l'IoT

Pour apporter une réponse aux questions scientifiques identifiées ci-dessus, il faut des compétences pointues en informatique, appliquées à différents domaines :

- *Réseaux de communications,*
- *Représentation de données,*
- *Systèmes distribués,*
- *Cryptologie,*
- *Traitement de données et protection de la vie privée,*
- *Sûreté, fiabilité et certification,*
- *Interaction Homme-machine,*
- *Contrôle,*
- *Sécurité,*
- *Architecture matérielle basse consommation, programmation et compilation,*
- *Optimisation globale de l'empreinte ressources.*

Les chercheurs d'Inria mènent des travaux scientifiques qui contribuent à faire progresser l'état des connaissances dans la plupart des domaines de recherche ainsi identifiés, qui sous-tendent l'IoT. Dans la deuxième partie de ce document, nous examinerons plus précisément les différents défis que doit relever la recherche sur l'IoT dans les différents domaines cités.

PARTIE 2

Domaines de Recherche pour l'IoT

Dans cette partie, nous explorons les différents domaines de recherche en informatique que l'IoT exploite, et au sein desquels des avancées sont nécessaires afin de répondre aux questions scientifiques clés que nous avons identifiées dans la partie précédente.

> Les chapitres ci-dessous (qui couvrent chacun de vastes domaines de recherche) peuvent être lus dans n'importe quel ordre, et se veulent autonomes. Par exemple, le lecteur peut choisir de lire le chapitre sur la Cryptologie *avant celui sur les* Réseaux de Communication. *Au sein de chaque chapitre, le document ne vise pas à couvrir de manière exhaustive les sujets de recherche connexes, mais plutôt à mettre en évidence la diversité des problèmes abordés.*

> Les lecteurs pressés (et ceux que des approfondissements scientifiques/ technologiques intéressent moins) peuvent se contenter de parcourir brièvement les titres des sections qui suivent, pour un aperçu rapide avant d'atteindre la conclusion du document..

2.1 Réseaux de communication pour l'IoT

Avec la montée en puissance de l'IoT et de la communication de machine à machine (M2M), d'importants bouleversements sont à prévoir dans la périphérie et le cœur du réseau.

Optimisation des protocoles du réseau d'accès IoT

En périphérie du réseau, la multitude d'appareils IoT a besoin d'une connectivité efficace et disponible pour rallier le réseau. Des études montrent que nous approchons déjà d'un stade où 75 % des connexions internet proviennent du segment grand public. Les protocoles d'accès sans fil posent des difficultés particulières.

Les appareils IoT basés sur microprocesseurs (ci-après désignés sous le terme « IoT haut de gamme ») sont demandeurs de plus en plus de débit, et remettront inévitablement en cause les standards des connexions sans fil. Les technologies annoncées pour l'accès sans fil haute puissance (mesurée en Watts, comme le Wi-Fi 6 ou la 5G) auront fort à faire pour répondre aux prochaines applications à forte intensité réseau, au nombre desquelles la réalité virtuelle, la réalité augmentée ou les véhicules autonomes totalement immersifs et améliorés par la technologie IoT. La conception d'optimisations et de nouveaux protocoles dans cet espace continue donc d'offrir des problèmes de recherche pertinents. La difficulté au niveau technologie radio consiste à s'approcher de la limite de capacité de transport du medium sans fil (imposée par les lois de la physique et la théorie de l'information), pour exploiter au mieux le spectre des fréquences radio disponibles. Sur le plan matériel, il faut concevoir et produire de nouvelles puces destinées à des communications radio optimisées, ce qui nécessite d'investir massivement en R&D.

Certaines techniques comme le MIMO massif semblent prometteuses sur le plan de la densité, mais nécessitent d'être étudiées plus avant pour atteindre leur plein potentiel. De nouvelles solutions algorithmiques attirent également l'attention, qui s'appuient sur des techniques d'apprentissage automatique (apprentissage par renforcement, apprentissage profond). La conception d'optimisations et de nouveaux protocoles d'accès pour les appareils IoT haut de gamme devrait donc devenir un domaine de recherche dynamique et stimulant.

↗ Au sein d'Inria, l'équipe-projet **MARACAS** associe à la théorie des communications et à la théorie de l'information des approches de traitement du signal, de théorie du contrôle ou de la théorie des jeux pour explorer de nouvelles technologies qui optimisent les communications dans les couches physiques (PHY) sans fil. Au-dessus des couches PHY sans fil, les équipes-projets **TRIBE** et **EVA** se consacrent à l'optimisation des accès multiples sur sans fil (par exemple accès aléatoire), en exploitant notamment des techniques d'apprentissage automatique.

L'exploration de techniques de communication complémentaires et radicalement différentes permet d'alléger la pression sur la ressource radio. De nouveaux travaux et de nouveaux défis ne manqueront pas de naître de la recherche sur des paradigmes alternatifs de communication comme les nano-communications ou les communications par la lumière visible (*visible light communications*, VLC). Ces technologies émergentes, à la frontière de la recherche en physique, méritent d'être explorées avec un esprit pionnier.

↗ Chez Inria, les équipes-projets **AGORA** et **FUN** étudient de nouvelles technologies de communication pour les réseaux d'accès alternatifs, utilisant les communications par la lumière visible (VLC), les communications RFID et des infrastructures hybrides, ainsi que les nouveaux services qu'elles pourraient fournir.

Étendre la couverture du réseau d'accès IoT

Les appareils IoT basse consommation (ci-après désignés sous le terme « IoT bas de gamme », qui utilisent des microcontrôleurs) s'appuient sur des technologies et des protocoles différents de ceux des appareils IoT haut de gamme pour les accès au réseau. Plutôt que de viser un débit maximum en supposant que les appareils disposent d'une puissance exprimée en Watts, les protocoles faible puissance ciblent une consommation d'énergie beaucoup plus basse, de l'ordre du milliwatt (voire moins) pour les appareils IoT de débit faible à moyen. Le domaine des technologies radio basse consommation pour les réseaux personnels/locaux (PAN/LAN) vise une **consommation d'énergie de plus en plus faible pour des distances de un à cent mètres.** L'objectif est d'être suffisamment économe pour ne nécessiter que l'énergie ambiante captable pour fonctionner. Quant aux technologies radio faible puissance pour les réseaux longue portée (IpWAN), elles cherchent à compenser la relative faiblesse du débit par des distances beaucoup plus importantes (dix kilomètres ou plus) à consommation d'énergie identique.

↗ Au sein d'Inria, l'équipe-projet **DIONYSOS** travaille à la conception de mécanismes de contrôle d'accès sans fil pour les grands réseaux NB-IoT.

L'une des difficultés majeures réside dans **l'extension de la couverture du réseau** à une myriade d'appareils IoT à la fois peu coûteux et économes en énergie. Il s'agit non seulement de gérer de façon efficace des réseaux d'accès surchargés, mais aussi de relever des défis sur les plans géographique et économique. Sur le plan géographique, la couverture doit s'étendre plus loin à l'intérieur des bâtiments, mais aussi plus loin vers l'extérieur, jusqu'aux zones les plus reculées. Sur le plan économique, la **technologie d'accès au réseau doit rester extrêmement abordable pour les appareils bas de gamme.**

Une portée de l'ordre du kilomètre ne suffit toutefois pas pour couvrir des appareils vraiment distants (comme dans le cas d'usage de l'agriculture intelligente dans lequel les capteurs peuvent être déployés dans des champs très vastes, loin des infrastructures urbaines) ou pour fournir un débit suffisant pour répondre aux exigences des applications. Par exemple, dans ce domaine, il reste à déterminer quelles sont les performances réalisables en matière de connectivité d'accès IoT utilisant les flottes de satellites déployés en orbite basse. D'autre part, les communications sans fil multibonds peuvent apporter une solution de remplacement, éventuellement en mettant à profit plusieurs technologies radio, mais elles demandent la création de nouveaux protocoles de routage adaptés à ces applications et environnements.

↗ Chez Inria, des équipes-projets comme **AGORA**, **FUN**, **MARACAS** et **TRIBE** conçoivent des optimisations de performance applicables au déploiement de capteurs sans fil basse consommation, et aux protocoles de routage et de diffusion de données multibonds dans ce contexte, à l'aide de méthodologies combinant modélisation théorique, simulation et expérience.

Un autre champ de la recherche s'efforce de compléter l'infrastructure fixe par une flotte de robots temporaires, en déployant des essaims de véhicules aériens sans pilote (aéronefs téléguidés, drones) ou de robots terrestres. On déploiera rapidement ces ressources complémentaires pour rétablir la connectivité après une panne, pour surveiller un événement ponctuel, pour explorer ou surveiller une zone inconnue/hostile ou simplement pour relever périodiquement les données contenues dans des dispositifs distants, que l'on ne pourrait atteindre par d'autres moyens. L'idée consistant à déployer des engins mobiles sans pilote (roulants ou volants) devient plus intéressante à mesure que les dispositifs commercialisés gagnent en maniabilité. Il reste cependant à trouver des solutions donnant à ces

appareils une autonomie totale, avec non seulement la contrainte d'une très faible consommation d'énergie, d'une faible puissance de calcul et d'une capacité de stockage limitée, mais également la nécessité d'assurer une bonne coordination, une communication et un partage des tâches efficace.

↗ Au sein d'Inria, des équipes-projets, dont **ACENTAURI**, **FUN** et **DANTE** conçoivent des algorithmes d'autodéploiement pour aéronefs téléguidés et robots terrestres, visant à optimiser le positionnement des robots et/ou à augmenter leur autonomie.

Optimisation des protocoles réseau pour le cœur et la périphérie

Le cœur du réseau doit actuellement transporter chaque mois des données dont le volume se mesure en exaoctets (milliards de milliards d'octets). Demain, avec la croissance rapide du trafic IoT (de machine à machine), il devra **transporter encore plus de données, avec une augmentation représentant des proportions considérables.**

Une difficulté majeure consiste donc à limiter le rythme de croissance du trafic IoT transitant par le cœur du réseau. Pour cette raison, un pan actif de la recherche explore **des architectures de protocole réseau alternatives pour une meilleure utilisation du stockage et du traitement des données dans le réseau,** aussi près que possible de l'origine des données (telles que le *Edge Computing*, ou encore *Information-Centric Networking*).

Une partie de la recherche dans ce domaine se focalise sur la latence. La latence moyenne dans la boucle entre la détection et l'action utilisant le réseau peut être de quelques millisecondes. Certaines applications IoT en temps réel (par exemple, l'Internet tactile, la téléchirurgie…) exigent cependant une latence strictement inférieure à quelques dizaines de millisecondes. Par conséquent, il y a un fossé à combler pour **répondre aux exigences combinées de la latence ultra-basse, du coût et de la complexité pour les applications IoT temps réel.**

↗ Chez Inria, l'équipe-projet **DIANA** mène des recherches sur la conception, la mise en œuvre et l'analyse de nouvelles architectures réseau, de services et de protocoles qui permettront la transparence des services et un meilleur contrôle des données utilisateur dans le contexte de centaines de milliards d'appareils mobiles connectés.

Normalisation des communications pour l'IoT

Par essence, la technologie IoT consiste à permettre à des systèmes embarqués, hétérogènes, de se connecter et d'interagir sur le réseau, potentiellement à grande échelle. Sans surprise, c'est donc un rôle central que jouent sdans l'espace IoT les organismes de normalisation , parmi lesquels :

- **IEEE,** qui travaille sur les protocoles de communication des couches physiques et de la couche lien, et en particulier le groupe de travail *IEEE 802.15*, qui développe les normes sans fil Bluetooth et *IEEE 802.15.4* par exemple ;
- **IETF,** qui travaille sur les protocoles de communication des couches réseau, transport et application, avec de nombreux groupes de travail chargés du développement de la pile de protocoles réseau sécurisé pour l'IoT basse consommation basée sur *6LoWPAN* et *IPv6* ;
- **W3C,** et en particulier le groupe de travail dédié au "*Web of things*", qui développe des programmes d'identification des ressources web pour l'interopérabilité sémantique entre les fournisseurs et les consommateurs de services IoT ;
- **3GPP** qui est le principal organisme de normalisation qui développe des protocoles de télécommunication cellulaire mobile (NB-IoT, 5G, etc.)..

↗ Chez Inria, des équipes-projets dont **EVA**, **PRIVATICS**, **TRIBE** et **WIMMICS** contribuent activement à la conception et à la normalisation de protocoles et de modèles de données au sein d'organismes de normalisation comme IETF et W3C.

Toutefois, l'écosystème des organismes de normalisation est en perpétuelle évolution et s 'étend au-delà de l'*IEEE*, de l'*IETF*, du *W3C* et du *3GPP*. Il est relativement malaisé de se repérer dans ce **paysage évolutif de normes qui se chevauchent et se concurrencent en partie.**

L'accès au réseau par radio utilise les normes *LoRa, Sigfox, NB-IoT, Bluetooth LE, ZigBee, Dash7, EnOcean, WirelessHART, DECT ULE, UWB*... l'accès au réseau par liaison filaire, les normes PLC, KNX, BACnet, CAN... De nombreuses normes de communication et de modèles de données sémantiques de haut niveau sont publiées par différents organismes et alliances, dont notamment : *OMA SpecWorks* (qui développe les normes *LWM2M*), *OPC Unified Architecture (OPC UA), OASIS* (qui développe les normes *MATT*), *DotDot (Zigbee Alliance), One Data Model (OneDM)*...

En raison du paysage fragmenté qu'offre l'IoT, mais aussi des spécifications des protocoles IoT, garantir l'interopérabilité et la sécurité entre les appareils de différents fournisseurs reste une gageure. **De nouvelles alliances se créent et se**

multiplient - qui développent des cadres d'interopérabilité et de certification pour différents secteurs d'activité – parmi lesquelles le *Thread Group, Connected Home over IP* (*CHIP*, récemment renommé *Matter*), *Zigbee Alliance, Open Connectivity Foundation (Alljoyn), WiSun*...

La consolidation du paysage des normes des technologies de communication utilisées par l'IoT fait partie des grands défis à relever. Avec l'avènement de l'IoT, des techniciens d'horizons et de cultures très différents sont amenés à travailler ensemble. Par exemple, des ingénieurs en matériel embarqué doivent maintenant collaborer avec les développeurs de protocoles Internet et de logiciels. Ces univers différents ont souvent des difficultés à se parler, ce qui révèle des problèmes de base, comme par exemple un manque criant de terminologie commune.

Même au sein d'une seule discipline technique, l'hétérogénéité extrême est un paramètre de l'IoT à prendre en compte, à différents niveaux : matériels, logiciels, protocoles réseau et normes technologiques... Il faut favoriser la convergence des technologies IoT vers une poignée de normes (dont certaines de fait), qui feront apparaître un terrain idéal favorisant des progrès plus rapides et une interopérabilité à grande échelle. Toutefois, la contrainte reste d'éviter les écueils d'une « monoculture ». En bref : le marché n'a pas encore accouché de la nécessaire consolidation des technologies IoT.

Protocoles réseau de bout en bout à basse consommation

LDes piles de protocoles inter-réseau, offrant polyvalence et faible consommation énergétique existent déjà. Le meilleur exemple en est probablement la pile de protocoles *IPv6* sur *6LoWPAN* et *CoAP*, normalisée par l'*IETF*. Dans la pratique, il subsiste toutefois à différents niveaux des problèmes de compatibilité. Par exemple, *IPv6* n'est souvent pas supporté nativement en périphérie de réseau, qui supporte uniquement *IPv4*. Un autre problème apparaît souvent entre les prestataires de services *cloud* et les vendeurs d'appareils IoT : ces derniers utilisent le protocole *CoAP* sur *UDP* au-dessus de la couche de transport, tandis que les premiers pensent HTTP sur TCP. La représentation des données utilisée est une autre source d'incompatibilité : typiquement, la sémantique ne correspond pas entre des données agrégées provenant de différents appareils IoT.

Des possibilités de contournement existent (par exemple, pour les incompatibilités citées ci-dessus, tunnel *Ipv6* sur *IPv4*, proxy CoAP-HTTP, traduction du code de données). Ces problèmes de compatibilité sont toutefois une barrière à l'entrée

importante (voire rédhibitoire) pour la plupart des non-spécialistes. Éliminer ces barrières reste donc un défi à relever. Supprimer ces obstacles permettra non seulement de réduire considérablement le coût d'entrée de l'IoT, mais aussi d'ouvrir la voie à la normalisation de paradigmes avancés de stockage et de calcul réseau pour l'IoT, dont l'informatique en périphérie de réseau a besoin, et de soulager la charge du réseau d'infrastructure.

Si l'on voit plus loin que ces solutions de contournement (proxy, passerelle…), le défi des nouvelles normes de protocoles réseau basse consommation est de parvenir à une large adoption, de bout en bout, tout au long du continuum *"Cloud-Edge-objet"*. Comme les ressources des appareils IoT basse consommation sont extrêmement contraintes, les solutions de bout en bout ne peuvent pas simplement ajouter une nouvelle couche de protocole sur l'existant. Par conséquent, la question de l'intégration des normes de protocoles réseau basse consommation est centrale.

En définitive, jusqu'où peut-on étendre le principe du réseau de bout en bout pour englober les appareils IoT basse consommation ?

2.2 Représentation de données pour l'IoT

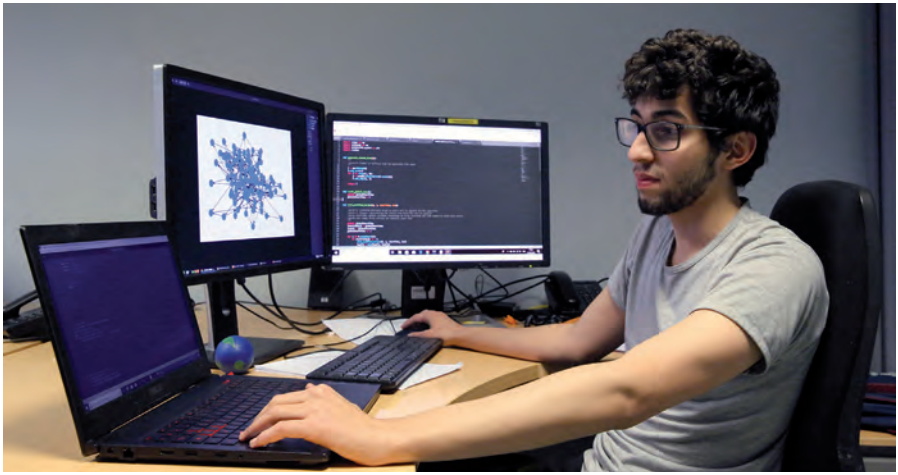
Nous connectons de plus en plus d'objets à Internet, qui deviennent visibles via différentes applications fonctionnant sur le réseau. Si ces objets produisent globalement de vastes quantités de données IoT brutes, ces données ne sont pas simples à exploiter en masse. Les appareils IoT (les « objets ») sont hétérogènes et utilisent souvent différents schémas de représentation des données, avec une sémantique variable pour les décrire. Cette réalité fragmente encore davantage l'IoT, au niveau de la couche applicative. On a besoin de nouvelles approches pour que les développeurs puissent créer des applications qui couvrent la diversité des objets et des technologies : pour relier les objets aux autres parties du système, il faut un cadre unifié.

Le Web des objets (Web of Things (WoT) wcherche à apporter un élément de réponse sur ce point. Il s'agit de s'appuyer sur le Web comme plate-forme d'application universelle pour les objets connectés. Littéralement, d'envisager **un Web du tout, fonctionnant sur n'importe quel support**. Cette perspective présente tout de même une difficulté, celle de faire évoluer les techniques classiques du Web pour tenir compte de la taille, de l'hétérogénéité et des spécificités des appareils et réseaux IoT.

Le Web a par exemple apporté des concepts centraux comme les URL, qui permettent d'identifier les appareils, les services et les acteurs. Ce mécanisme de référencement par défaut permet d'obtenir des descriptions riches pour les nouveaux identifiants découverts. Pour que le WoT devienne réalité, les modèles standard basés sur le Web doivent atteindre de nouveaux niveaux de flexibilité, répondant aux besoins :

- des vendeurs, pour décrire les caractéristiques de leurs produits et services ;
- des fournisseurs de plates-formes et d'applications, pour présenter ces caractéristiques ;
- des utilisateurs, pour exprimer leurs objectifs, leurs demandes, etc..

Cette latitude accordée dans la description est essentielle également pour que les fabricants puissent différencier leurs produits de ceux de leurs concurrents et offrir une gamme de modèles possédant des caractéristiques distinctives, tout en préservant leur interopérabilité à l'échelle du Web. Dans certains cas, ces modèles génériques doivent appliquer des normes partagées (concernant par exemple les types de données de base, les unités physiques) et fournir des représentations



Apprentissage automatique distribué pour les applications d'IoT pour le pilotage de la création et de l'évolution de réseaux complexes. © Inria / Photo L. Jacq.

neutres sur le plan du langage de programmation, des représentations qui rendent possible l'interopérabilité entre plates-formes et domaines. Cependant, ils doivent également reconnaître les extensions de modèles spécifiques aux applications et aux domaines, comme un typage et des définitions plus précis correspondant à des structures de données plus complexes destinées à des utilisations et à des scénarios spécifiques.

Les langages du Web, et en particulier le ***Linked Data*** et le **Web sémantique**, peuvent assurer une interopérabilité riche et des formats standard pour les descriptions d'objets, d'opérations, d'entrées, de sorties, etc. Le WoT crée un besoin de description. Une question ouverte spécifique au Web sémantique se pose : comment formaliser des vocabulaires multiplates-formes et multidomaines ? La difficulté consiste à fournir des langages et des vocabulaires offrant l'expressivité adéquate pour représenter formellement les « jumeaux numériques » des objets. Dans le WoT, les objets sont considérés comme des ressources logicielles identifiées sur le Web, dont il faut décrire les caractéristiques, les opérations et les événements et les relier pour favoriser la découverte, l'interopérabilité et la composition. Les modèles à concevoir doivent non seulement nous permettre de décrire des objets très hétérogènes (leurs besoins, leurs capacités et leurs caractéristiques), mais aussi leur plate-forme et leur contexte cyberphysique.

Puisque nous devons représenter, publier, interroger et valider ces métadonnées et les données échangées sur le WoT et en déduire des éléments, nous devons concevoir et normaliser des modèles abstraits, des syntaxes concrètes et des numérotations efficaces, des langages de traitement applicables dans un contexte dynamique aux ressources limitées, etc. Les approches du type *linked data* classiques sur le Web sémantique peuvent apporter des solutions (par exemple, les langages ontologiques), mais elles se heurtent également à des difficultés spécifiques au WoT.

Un défi majeur pour les modèles du Web sémantique réside dans la capacité à gérer le caractère dynamique du WoT et ses flux de données (par exemple, le produit des capteurs), et la reconfiguration rapide lorsque des objets sont connectés ou retirés. La nécessité de saisir le contexte de l'application et de s'y adapter en termes de portée, de distribution, de limitations, de confidentialité, de profils d'utilisateurs, etc. pose d'autres problèmes. Les modèles de données standard doivent également prendre en charge des langages de script web répondant aux besoins de l'interaction entre objets, indépendamment de la plateforme, pour les applications et la gestion. Les langages utilisés pour les requêtes et la validation des données liées peuvent servir à accéder aux descriptions et les valider par rapport aux contraintes.

Les représentations et le raisonnement basés sur l'ontologie peuvent contribuer à la composabilité et à l'interopérabilité avec les langages pivots et les transformations. Les langages pivots et les mécanismes de transformation fournis par le Web sémantique peuvent également jeter des ponts entre différents protocoles sous-jacents (par exemple, HTTP, *WebSockets*, *CoAP*, MQTT) et fournir des moyens uniformes pour échanger des messages avec les objets et les services, comme des liaisons déclaratives. Des dispositifs de sécurité peuvent aussi être exprimés et échangés entre systèmes hétérogènes.

↗ Au sein d'Inria, l'équipe-projet **WIMMICS** travaille sur l'intégration d'agents autonomes dans le *Web of things* (WoT), en s'appuyant sur les langages du Web sémantique et sur les principes des données liées (*Linked Data*) pour faire le lien entre l'architecture web et l'architecture des systèmes multiagents. L'objectif est de proposer un environnement basé sur un standard ouvert pour déployer des agents et des comportements intelligents dans les systèmes IoT, dans les scénarios d'automatisation, par exemple dans une usine.

2.3 Systèmes distribués pour le continuum *Cloud-Edge-Objet*

La conception, le développement et l'exécution d'applications dans l'IoT nécessitent de maîtriser et de gérer sa complexité en termes de distribution, d'hétérogénéité, de dynamicité et d'adaptation.

Intergiciels pour l'IoT

Les intergiciels permettent habituellement de traiter des problématiques de dynamicité et d'hétérogénéité de façon transparente pour les systèmes distribués, soit sous la forme d'une couche logicielle opérant sur chaque appareil, soit via une entité logicielle présente quelque part dans le réseau et jouant le rôle d'intermédiaire. La conception des intergiciels est toutefois confrontée à de nouveaux défis particuliers lorsqu'il s'agit de prendre en charge des applications dans le domaine de l'IoT.

Le principal défi consiste à s'adapter au degré élevé d'incertitude caractéristique de l'environnement d'exécution des objets connectés, qui contraste avec le processus typique d'ingénierie logicielle, dans lequel un système est totalement finalisé pendant sa phase de conception. Le contexte de l'IoT est en constante évolution, et la complexité des changements (auxquels le système IoT doit s'adapter) est telle qu'elle ne peut être traitée au moment de la conception du système. Du fait de leur composition et de leur fonctionnement automatisés, dynamiques et dépendants de l'environnement, des systèmes IoT peuvent faire leur apparition de façon inattendue. Les systèmes et leurs propriétés ne prennent leur forme complète qu'au moment de l'exécution, et ils évoluent généralement par la suite, nécessitant des niveaux d'interopérabilité imprévus.

Le *crowdsensing* de masse est l'une des principales applications des intergiciels IoT, par exemple pour des mesures dans l'environnement urbain. La conception de nouveaux algorithmes et protocoles capables à la fois de prendre en compte efficacement l'utilisation massive de *smartphones* et d'autres objets connectés, et de gérer la participation dynamique à grande échelle des utilisateurs au sein

de l'IoT représente un autre défi pour ces intergiciels. Outre le *sensing* physique, dans lequel le capteur d'un appareil signale en tâche de fond les phénomènes détectés, le social *sensing* entre en jeu lorsque l'utilisateur est conscient et participe activement au *sensing*.

Un autre défi découle ici du manque de précision des capteurs et de l'absence de contrôle des conditions dans lesquelles se déroule le *sensing* participatif. Il est particulièrement difficile de faire du *crowdsensing* opportuniste un moyen fiable pour étudier les phénomènes environnementaux. La conception de mécanismes de coordination spontanée et décentralisée entre les capteurs mobiles fait l'objet de travaux de recherche.

↗ Chez Inria, des équipes-projets dont **MIMOVE** et **SPIRALS** étudient des solutions intergicielles pour l'IoT. DeXMS, par exemple, est un intergiciel développé par Inria dans le but d'apporter aux systèmes IoT émergents des propriétés dynamiques en matière d'interopérabilité, de composition et de programmation, tout en s'appuyant sur des ressources informatiques situées à la périphérie du réseau. D'autres plates-formes intergicielles conçues par Inria, telles que APISENSE ou *SenseTogether*, ciblent des applications du type mobile *crowd sensing*. Ces intergiciels visent à améliorer la qualité des données fournies par les objets connectés en exploitant la connaissance du contexte et les ressources à la périphérie du réseau, y compris les capteurs participatifs mobiles eux-mêmes.

Plates-formes de test pour le continuum *Cloud-Edge*-Objet

L'IoT est composé de milliards d'appareils connectés offrant des capacités de calcul restreintes, mais qui peuvent collaborer sur le réseau. Cette collaboration s'opère soit entre objets connectés, soit *via* des machines situées à proximité, qui offrent des capacités de calcul moyennes (on parle également d' *"edge"* , de *"fog computing"*), soit au moyen de systèmes plus éloignés offrant de grandes capacités de calcul (*"cloud computing"* ou « informatique en nuage »). Cette collaboration nécessite toutefois l'existence d'un continuum *Cloud-Edge/Fog*-Objet permettant de distribuer les calculs de manière adéquate suivant les exigences (qui sont elles-mêmes susceptibles d'évoluer dans le temps).

↗ Inria dirige le développement de grandes infrastructures de test pour la recherche expérimentale sur le continuum *Cloud-Edge*-Objet, notamment SILECS, une plate-forme de test en accès libre permettant aux chercheurs de créer et de déployer de bout en bout une pile logicielle complète (la partie système et les protocoles) depuis les plus petits objets connectés jusqu'aux grands centres de données. L'objectif est de permettre la capture globale et fine des événements, depuis le phénomène physique observé (capteurs/actionneurs) jusqu'au traitement et au stockage des données, en passant par le déploiement dynamique de services de calcul en périphérie, et les transmissions radio.

Orchestrer les ressources *Cloud-Edge*-Objets

Les concepteurs de systèmes cyberphysiques sont chargés de programmer le continuum *Cloud-Edge/Fog*-Objets. Jusqu'à présent, ils utilisaient des techniques et modèles très différents pour programmer et gérer ces différentes catégories de machines (*Cloud, Edge*, objets), rendant ainsi difficiles la compréhension et l'évaluation du système dans son ensemble. L'orchestration dynamique des ressources *Cloud-Edge/Fog*-Objets est en effet un défi. L'évaluation des caractéristiques globales de sécurité et de confidentialité d'un système cyberphysique fait partie des défis connexes. Dans ce domaine, une piste de recherche considérée est la conception d'une syntaxe unifiée, susceptible d'être utilisée pour programmer **tous les composants** du système cyberphysique et de permettre une caractérisation globale du système. Cette approche vise à traiter et à faire appliquer des mesures de sécurité sur l'ensemble du continuum (et non pas composant par composant) et à simplifier la programmation du comportement temporel non trivial des objets connectés, qui allie des activités synchrones et asynchrones.

↗ L'équipe-projet **INDES** d'Inria travaille à la création de langages de programmation multicouche sécurisés pour l'IoT, exprimant une syntaxe du continuum, depuis les microcontrôleurs jusqu'au *Cloud*. Les dialectes JavaScript Hop.js and HipHop.js. sont des exemples de telles syntaxes qui permettent d'utiliser le même code pour programmer les serveurs, les clients et les objets connectés, simplifiant ainsi la conception du système IoT et assurant l'application générale de mesures de sécurité.

DevOps pour les systèmes cyberphysiques

Des chaînes d'outils puissantes sont nécessaires pour raccourcir le délai entre le lancement d'un changement dans un système distribué et la mise en production de ce changement, tout en assurant un niveau de qualité élevée. *DevOps* est à la fois un environnement technique et un domaine de recherche qui utilise et développe de telles chaînes d'outils pour les logiciels, englobant toutes les étapes (codage, *build*, test, conditionnement, déploiement, configuration et supervision) pendant la durée de vie du système. Dans la pratique, *DevOps* est nécessaire pour permettre le développement et la maintenance agiles qui sont attendus des logiciels modernes à l'ère d'Internet.

Les chaînes d'outils *DevOps* traditionnelles ne sont généralement pas applicables aux petits objets connectés en raison des contraintes liées aux réseaux et ressources embarquées de faible puissance. Avec la présence d'objets très contraints au sein du continuum (*Cloud-Edge/Fog*-Objets), un enjeu réside donc dans la conception et la mise en œuvre de nouvelles chaînes d'outils *DevOps* plus complètes, qui étendent leur champ d'application aux objets connectés plus petits et de faible puissance. Une problématique connexe est la nécessité d'évaluer et de garantir les caractéristiques de sécurité de ces chaînes d'outils étendues.

Corriger des bugs dans des objets IoT opérés à distance avec Pharo. © Inria / Photo Raphaël de Bengy.

↗ L'équipe-projet **TRIBE** d'Inria travaille à la conception d'approches *DevOps* sécurisées qui peuvent s'étendre aux objets connectés de faible puissance, applicables non seulement aux machines connectées utilisant des microprocesseurs, mais aussi à celles comprenant des microcontrôleurs.

Placement optimal du calcul dans l'IoT post-*Cloud*

Le traitement dans le *Cloud* des données provenant d'objets connectés peut être complété (ou entièrement contourné) en exploitant une puissance de calcul plus proche de l'origine des données brutes. Les systèmes géodistribués intermédiaires (*Edge Computing*) peuvent contribuer à la puissance de calcul, de même que, dans une certaine mesure, les objets connectés bas de gamme eux-mêmes. L'IoT post-*Cloud* offre potentiellement des capacités de (pré-)traitement des données à n'importe quel endroit du réseau, de bout en bout. Face à une telle possibilité, un nouvel enjeu consiste à identifier et à mettre en œuvre des stratégies pour le placement optimal des services IoT – calcul ou prétraitement des données – le long du continuum (*Cloud-eEge/Fog*-objets). Un autre défi se pose alors, celui de l'automatisation de la migration des services le long du continuum, afin d'adapter dynamiquement le système cyberphysique à des exigences haut niveau (qui évoluent régulièrement). Des communautés de chercheurs telles que COIN, par exemple, explorent des architectures de réseau alternatives tels *software-defined networking (SDN)* ou encore *information-centric networking (ICN)* ainsi que la virtualisation des fonctions réseau (NFV), qui ont pour but de relever ce défi.

↗ L'équipe-projet **STACK** d'Inria travaille à la conception de mécanismes systémiques et d'abstractions logicielles pour gérer et utiliser des infrastructures d'informatique utilitaire nouvelle génération, associant *Cloud, Fog* et *Edge*. Ces techniques visent à gérer efficacement le cycle de vie des applications fonctionnant sur le continuum *Cloud*-IoT. Elles prennent en considération les coûts et des aspects transversaux tels que la consommation d'énergie, la latence de lancement des applications, les contraintes de bande passante et la sécurité.

De nouvelles architectures d'accès sans fil apparaissent, entraînant une *"softwarisation"* accrue de l'infrastructure réseau (la 5G cellulaire, et au-delà, en est un exemple notable). La tendance est à l'abandon des points d'accès matériels coûteux et dédiés (propriétaires) au profit d'antennes « simples » associées à des logiciels *backend* exécutés sur des serveurs génériques bon marché dans le *Cloud*

(ou en périphérie du réseau). Ces réseaux à définition logicielle (SDN) augmentent considérablement la flexibilité de l'infrastructure d'accès sans fil, et pourraient bouleverser les modèles économiques constructeur/opérateurs. Ces architectures peuvent non seulement permettre un « découpage » plus pointu des ressources d'accès au réseau, mais aussi fournir, de façon dynamique, de nouvelles capacités de calcul spécifiques aux utilisateurs. De nouveaux compromis sont envisageables en matière de coût, de latence, de fiabilité, etc.

Un autre aspect associé est la mobilité des utilisateurs d'objets connectés, à la fois source de défis et d'opportunités. En exploitant les modèles spatio-temporels de mobilité des utilisateurs, le système pourrait servir de base à des stratégies d'allocation des ressources plus précises et plus agiles. Dans ce domaine, les enjeux consistent, d'une part, à caractériser (et à anticiper) la mobilité des utilisateurs d'objets connectés, d'autre part à concevoir des schémas dynamiques de déchargement et de placement du calcul permettant d'optimiser les performances selon ces prévisions.

↗ L'équipe-projet **TRIBE** d'Inria travaille à la conception de nouvelles stratégies de déchargement et de placement du calcul pour l'IoT qui exploitent la mobilité des utilisateurs, et à l'évaluation de l'impact du transfert des tâches à la périphérie du réseau en termes de consommation d'énergie et de latence, dans des contextes d'IoT mobile..

Maintenance des logiciels distribués pour l'IoT

Alors que les objets connectés sont toujours plus nombreux à être déployés, une grande partie d'entre eux n'intégre pas de mécanisme sécurisé pour la mise à jour de leurs logiciels. Et si certains embarquent un mécanisme de ce type, il reste souvent inutilisé en fin de compte, pour des raisons qui n'ont rien à voir avec la technique : par exemple un manque d'incitations de la part du fournisseur (ou parce que celui-ci a fait faillite entre temps). Des vulnérabilités critiques en matière de sécurité ne sont donc jamais corrigées, et les objets connectés deviennent une épée de Damoclès, comme l'ont montré de récentes cyberattaques réalisées à grande échelle. À titre d'exemple, de nombreux éléments du *botnet* IoT *Mirai* (ainsi que d'autres cibles potentielles de ce botnet) fonctionnent encore sans correctif à ce jour, bien que Mirai ait été découvert il y a plusieurs années, que des correctifs aient été développés depuis, et que ces machines ne soient pas parmi les plus limitées en ressources !

Sur le plan juridique, une problématique qui émerge concerne le devoir de prudence : qu'est-ce ce devoir rend obligatoire pour l'objet connecté avant sa mise en service ? Qu'en est-il pour les objets connectés déjà déployés (matériel en service, et hérités) ?

Sur le plan technique, l'enjeu important est de rendre possibles (et d'automatiser) les mises à jour légitimes des logiciels s'exécutant sur les objets connectés. Sur les objets connectés de faible puissance, la difficulté est amplifiée par des contraintes de ressources strictes en termes de débit réseau, d'énergie et de budget mémoire. D'un autre côté, imposer une vérification de la légitimité des mises à jour logicielles peut conduire à ce que l'on appelle l'informatique déloyale qui, paradoxalement, empêche les mises à jour logicielles qui seraient nécessaires.

Les recherches antérieures menées dans ce domaine sur les objets connectés de faible puissance se sont principalement concentrées sur des cas d'utilisation simples où les mises à jour ciblaient des logiciels à code binaire unique et à acteur unique. Mais, étant donné que les logiciels IoT évoluent et se complexifient, cette simplification n'est plus viable : les logiciels IoT imitent de plus en plus les logiciels internet dans le sens où ils sont un amalgame de composants disparates, développés, maintenus et mis à jour par différents acteurs. Aujourd'hui, dans une entreprise de taille moyenne, moins de 5 % des logiciels utilisés sont développés en interne, plus de 95 % provenant de tiers et/ou sont des logiciels libres.

Un défi à relever consiste donc à sécuriser efficacement la mise à jour des logiciels multiacteurs sur les objets connectés bas de gamme, sachant que ces différents acteurs ne jouissent que d'une confiance mutuelle limitée. Plusieurs considérations entrent en jeu, qui mènent à combiner les résultats issus de plusieurs domaines de recherche dont :

- de nouveaux mécanismes de systèmes profondément embarqués pour héberger et isoler mutuellement différents modules logiciels ;
- la création et la sécurisation de la chaîne d'approvisionnement des logiciels IoT ;
- des mécanismes de réseaux IoT à basse consommation énergétique performants pour le transport des mises à jour.

En pratique, la maintenabilité implique également la capacité à superviser et à gérer à distance, via le réseau, les objets connectés tout au long de leur durée de vie. En conséquence, un domaine de recherche actif est le développement de protocoles à faible consommation énergétique et de modèles de données de gestion adaptés aux objets connectés par radio basse puissance. Un pan de recherche associé concerne l'instrumentation efficace d'objets connectés, pour supervision à distance *via* le réseau, avec des bribes de code et de débuggage bas niveau, qui sont insérées et supprimées à la demande, en parallèle de l'exécution normale. D'autres travaux sont menés dans le domaine des logiciels adaptatifs : leur but est de passer d'un logiciel simplement adaptatif à un logiciel auto-adaptatif. En effet, une automatisation avancée pourrait permettre de mettre au point des logiciels autoréparables.

↗ L'équipe-projet **SPIRALS** d'Inria étudie des systèmes auto-adaptatifs visant à introduire un niveau d'automatisation accru les systèmes logiciels, en ciblant principalement les propriétés d'autoréparation et d'auto-optimisation.

Un dernier point, et non des moindres, concerne la maintenabilité des objets connectés, qui peut être fortement améliorée par une conception coévolutive des logiciels avec le matériel IoT. Du point de vue de la maintenance par exemple, l'implémentation logicielle d'une fonction cryptographique est à privilégier par rapport à une implémentation *hardware* (car cette dernière n'est pas modifiable par la suite pour corriger des bugs, ou pour mettre en œuvre une nouvelle règlementation). Un domaine de recherche associé vise donc conjointement à améliorer les performances d'implémentations logicielles pour certaines fonctionnalités critiques et à concevoir des accélérateurs matériels génériques en support.

QUELQUES MOTS SUR LES MODÈLES ÉCONOMIQUES.

Jusqu'à présent, la plupart des modèles économiques privilégient les gains financiers liés au déploiement et/ou à l'exploitation de l'IoT. On en sait moins sur les profits réalisables dans la maintenance des objets connectés, qui constitue pourtant un aspect crucial.

Une fois les objets connectés déployés et mis en service, il est indispensable d'instaurer des relations moins féodales entre utilisateurs et fournisseurs de l'IoT. Les mises à jour logicielles des objets connectés, en particulier, doivent être facilitées de manière générale, y compris dans les cas où le fabricant d'origine ne les fournit pas (parce qu'il a été racheté, a fait faillite ou a procédé à un retrait forcé du produit, par exemple). Au-delà des obstacles techniques, les barrières juridiques (comme la rupture du contrat/de la garantie) rendent souvent les mises à jour logicielles difficiles, voire impossibles. Ce phénomène ressemble au modèle dit du « jardin clos », qui ne permet aux utilisateurs que d'ajouter des composants matériels autorisés, ou d'acheter des services de réparation auprès de revendeurs agréés.

Les fournisseurs d'objets connectés, par exemple, qui sont tenus de certifier leurs produits IoT pour des raisons de sécurité, mettent rarement à jour les logiciels présents dans ces produits. Les utilisateurs, *a contrario*, peuvent souhaiter mettre à jour plus fréquemment le logiciel, notamment pour obtenir plus de fonctionnalités. Des rapports ont montré comment de telles tensions ont déjà entraîné des situations regrettables, dans lesquelles les utilisateurs ont en fin de compte recours à des logiciels IoT piratés, ce qui ne fait que compliquer les choses.

2.4 La cryptologie appliquée aux objets connectés « bas de gamme »

La cryptographie fournit les protocoles fondamentaux et les algorithmes de base (les « primitives ») servant à l'authentification, à l'identification et au cryptage sur lesquels reposent tous les systèmes sécurisés.

Les cryptographes ont des décennies d'expérience dans la conception et l'analyse de cryptosystèmes et de protocoles performants, d'une part pour des appareils relativement puissants (tels que les PC, les serveurs ou les *smartphones*), d'autre part pour des dispositifs plus limités comme les cartes à puce. L'essor de l'IoT, caractérisé par l'omniprésence d'appareils interconnectés de faible puissance, constitue un nouveau défi passionnant pour les cryptographes, qui doivent prendre en compte simultanément les exigences applicatives du paradigme PC et les contraintes physiques sévères liées aux appareils bas de gamme. En d'autres termes, nous savons comment assurer un certain niveau de sécurité pour les microcontrôleurs des cartes à puce, mais celles-ci n'ont jamais été conçues pour être connectées à Internet. Et nous savons comment assurer la sécurité Internet de puissants processeurs, mais pas avec un budget basse consommation strict. L'enjeu pour les cryptographes est de développer des primitives puissantes adaptées aux contraintes et exigences spécifiques du paradigme IoT.

Des primitives cryptographiques pour des communications IoT sécurisées

Des primitives cryptographiques extrêmement performantes et hautement sécurisées ont été normalisées et largement déployées dans des suites de protocoles tels que TLS (Transport Layer Security) pour assurer la sécurité des communications sur Internet. Toutefois, ces algorithmes ont traditionnellement été développés et optimisés pour des plates-formes plus puissantes, depuis les serveurs et PC jusqu'aux smartphones. Lorsque nous passons à des objets connectés bas de gamme plus limités, les contraintes de ressources se resserrent au point que les primitives classiques sont souvent trop coûteuses pour l'appareil concerné. L'enjeu fondamental en la matière est donc le développement, l'optimisation et l'adoption de primitives cryptographiques alternatives constituant des briques de base adaptées pour des communications IoT de faible puissance..

Primitives cryptographiques symétriques et asymétriques

Les primitives cryptographiques sont réparties en deux catégories de base, les primitives symétriques et les primitives asymétriques, selon leur fonction et leur application. Le cryptage et l'authentification des données sont, par exemple, des primitives symétriques, alors que l'échange de clés et les signatures sont des primitives asymétriques. De manière générale, les primitives symétriques ont un débit beaucoup plus élevé et une consommation de ressources plus faible. Les primitives asymétriques, quant à elles, offrent des fonctionnalités essentielles (telles que les signatures numériques), que la cryptographie symétrique est absolument incapable de fournir. Mais comparativement, ces primitives asymétriques nécessitent inévitablement des clés d'une taille plus élevée, une structure interne plus complexe et des calculs plus intensifs, qui impliquent certaines exigences en termes de temps de calcul, de mémoire et de batterie. En pratique, sécuriser les communications dans l'IoT requiert à la fois des primitives symétriques et des primitives asymétriques optimisées.

Primitives cryptographiques préquantiques et postquantiques

Avec l'essor de l'informatique quantique, il convient de faire une seconde distinction importante, cette fois entre les primitives pré et postquantiques. Cette distinction repose sur le changement de modèle d'attaque : si un cryptosystème est conçu pour résister aux attaques d'adversaires équipés aussi bien de calculateurs quantiques que d'ordinateurs classiques, il est alors qualifié de postquantique. La construction d'ordinateurs quantiques suffisamment puissants pour attaquer les cryptosystèmes modernes est un défi majeur pour les physiciens et les ingénieurs. Bien que nous ne puissions que spéculer sur le moment où ils y parviendront, ou sur le fait qu'ils réussiront, nous devons, quoi qu'il arrive, préparer les objets connectés à un avenir quantique. Nous ne pouvons tout simplement pas faire reposer la sécurité future sur la faiblesse de la science.

Cryptographie symétrique optimisée pour l'IoT

Les primitives symétriques nécessitent une clé secrète partagée par les parties communicantes. Parmi les exemples significatifs, citons les algorithmes de chiffrement des données, tels que *ChaCha20* et *AES* (adopté par le NIST et devenu *de facto* le standard international de référence), et d'authentification des messages comme *HMAC* et *Poly1305*. Les fonctions de hachage (telles que *SHA-3*), bien que n'utilisant généralement pas de clé, font également partie de la famille des primitives cryptographiques symétriques.

Si nous considérons maintenant les objets connectés bas de gamme, nous entrons dans le monde de la cryptographie légère. Elle vise à fournir des primitives symétriques efficaces avec des empreintes ressources extrêmement réduites, mais souvent assorties de niveaux de sécurité sensiblement inférieurs. La cryptographie légère intéresse l'IoT pour deux raisons : d'abord parce qu'elle permet d'effectuer des opérations cryptographiques dans des appareils ayant des ressources très limitées, ensuite parce qu'elle est requise pour les appareils bas de gamme communicants. Le NIST (l'influent organisme de normalisation américain) a lancé un processus de normalisation concurrentiel pour développer de nouvelles primitives légères. Son résultat aura un impact important sur la cryptographie symétrique dans l'espace IoT.

Dans la conception des systèmes cryptographiques pour les objets connectés bas de gamme, l'un des enjeux est d'échapper à la « double peine » subie sur les microcontrôleurs par rapport à l'environnement PC/smartphone : non seulement les processeurs (CPU) sont moins puissants et plus lents (pénalité de performance 1), mais les accélérateurs matériels font parfois défaut, ce qui oblige à adopter des approches purement logicielles (pénalité de performance 2). À titre d'exemple, certains microcontrôleurs ne disposent pas du support matériel pour l'AES (la norme NIST), qui est considéré comme incontournable dans le monde des PC. Au lieu d'une implémentation logicielle du cryptage AES imitant l'environnement PC, des primitives cryptographiques symétriques alternatives devraient être conçues pour être utilisées sur les objets connectés bas de gamme. L'expérience montre, par exemple, que le passage de l'AES à un schéma de chiffrement logiciel de type *ChaCha20* peut améliorer les performances de 30 % pour certaines applications.

↗ L'équipe-projet **COSMIQ** d'Inria travaille à la conception et à l'analyse de primitives symétriques légères. Les chercheurs de **COSMIQ** participent aux propositions faites dans le cadre du processus de normalisation de la cryptographie légère du NIST.

Cryptographie symétrique pour l'IoT postquantique

La cryptographie symétrique postquantique a principalement été étudiée en réponse à des menaces. C'est le cas de l'algorithme de Grover, qui, expliqué (très) grossièrement, permet de rechercher des ensembles de clés possibles en un temps proportionnel à la racine carrée du nombre de clés (alors que les ordinateurs classiques nécessitent un temps linéairement proportionnel au

nombre de clés). Il est communément admis que, pour de nombreuses primitives cryptographiques symétriques de base, il convient de doubler la longueur des clés pour assurer la sécurité postquantique. La problématique est, bien entendu, beaucoup plus subtile, d'autant plus si l'on prend en considération des systèmes et des opérations symétriques plus complexes. Le véritable niveau de sécurité postquantique des primitives symétriques existantes fait l'objet de recherches actives. Mais, même si le simple fait de doubler la longueur des clés suffisait, cela aurait un impact important sur la sécurité des objets connectés : outre le fait que cette solution doublerait l'espace requis pour les clés, des facteurs variables viendraient dégrader le débit et la consommation de ressources des algorithmes. Par exemple, le passage de la primitive cryptographique *SHAKE128* à la primitive correspondante avec une longueur de clé doublée (*SHAKE256*) n'aurait aucun impact sur la capacité mémoire requise, mais pourrait entraîner une diminution du débit d'au moins 20 %.

↗ L'équipe-projet **COSMIQ** d'Inria travaille sur la sécurité des cryptosystèmes symétriques postquantiques.

Scuba : une chaine d'outils pour la sécurité des objets connectés. © Inria / Photo D. Betzinger.

Cryptographie asymétrique optimisée pour l'IoT

Contrairement aux primitives symétriques, les primitives asymétriques reposent sur le fait que chaque partie conserve une « clé privée (ou secrète) » qui n'est jamais révélée, et diffuse une « clé publique » correspondante aux autres parties. Exemple : Alice signe un message avec sa clé privée ; plus tard, après avoir reçu le message, Bob peut vérifier la signature d'Alice en utilisant la clé publique de celle-ci. Cette asymétrie, illustrée par la distinction entre clés privées et clés publiques, rend possible un grand nombre de nouvelles primitives qui ne peuvent pas exister dans le modèle symétrique, notamment des signatures, mais aussi l'échange de clés Diffie-Hellman, élément fondamental pour créer des clés secrètes partagées et permettre une communication sécurisée par chiffrement symétrique.

Les clés publique et privée sont étroitement liées : en substance, la clé publique représente un problème mathématique (comme un logarithme discret à courbe elliptique), et la clé privée représente la solution à ce problème. Le problème est choisi de manière telle que sa résolution sur le plan informatique s'avère infaisable (ou loin d'en valoir la peine). De manière générale, donc, utiliser des cryptosystèmes asymétriques revient à se livrer à des calculs complexes dans des structures mathématiques, ce qui induit un coût élevé en termes de mémoire et d'énergie. Un coût bien souvent trop conséquent pour des appareils IoT bas de gamme. L'amélioration de la performance et de l'applicabilité des cryptosystèmes asymétriques dans l'espace IoT est donc un domaine de recherche actif.

Dans l'environnement préquantique, la cryptographie à clé publique pour l'IoT est dominée par la cryptographie sur les courbes elliptiques (ECC). L'avantage crucial de l'ECC est la taille particulièrement réduite de ses clés : 32 octets suffisent pour stocker une clé ECC haute sécurité, et une signature ECC haute sécurité tient dans 64 octets. Toutefois, l'utilisation de l'ECC implique d'effectuer un grand nombre de calculs (modulo) sur des entiers de 32 octets (et non de 32 bits !). Cela ne pose aucun problème dans le domaine des PC, et l'ECC est désormais universellement utilisée pour l'échange de clés et les signatures sur Internet. Toutefois, dans le domaine des objets connectés de faible puissance, ces calculs basés sur la théorie des nombres nécessitent une empreinte mémoire non négligeable, grèvent considérablement les réserves d'énergie et ralentissent le temps d'exécution en entraînant des latences. Un domaine de recherche vise donc à adapter les protocoles ECC et à développer de nouveaux algorithmes ECC pour faire plus avec moins, c'est-à-dire maintenir une sécurité élevée pour les objets connectés bas de gamme, tout en réduisant considérablement les coûts liés aux temps d'exécution.

↗ L'équipe-projet **GRACE** d'Inria travaille sur la sécurité side-channel des microcontrôleurs, et sur des primitives cryptographiques efficaces à clé publique (asymétriques), dont certaines ciblent les objets connectés, telles que le mécanisme de signature *qDSA*.

Cryptographie asymétrique pour l'IoT postquantique

Si l'on envisage l'avenir postquantique, la cryptographie asymétrique est confrontée à un problème majeur : l'existence d'ordinateurs quantiques de taille suffisamment conséquente exécutant l'algorithme de Shor anéantirait la sécurité de quasiment toutes les primitives asymétriques actuellement déployées. Un domaine de recherche vise donc à développer et à étudier d'autres systèmes de cryptographie asymétrique résistant aux attaques quantiques, qui mettent en œuvre une grande diversité d'approches, notamment des systèmes à base de treillis (lattice), de hachage, de code et d'isogénie.

Le NIST a lancé un processus international pluriannuel afin de sélectionner des algorithmes candidats pour les signatures postquantiques et l'encapsulation de clé (remplaçant essentiellement la méthode Diffie-Hellman préquantique). Plusieurs algorithmes candidats finalistes qui sont à l'étude réduisent déjà la taille requise des clés publiques, les signatures et les coûts de calcul. Ces algorithmes n'ont toutefois pas été conçus pour des applications IoT, et leur performance dans ce domaine reste à explorer. Le développement de mécanismes d'établissement de clé et de signatures postquantiques performants, pratiques et éprouvés pour l'espace IoT, est une source de problématiques absolument passionnante pour les chercheurs en cryptographie.

↗ Les équipes-projets **ARIC**, **COSMIQ** et **GRACE** d'Inria travaillent à la conception, à l'analyse et à l'implémentation efficace de cryptosystèmes asymétriques postquantiques. Les chercheurs d'Inria participent également aux propositions soumises au NIST pour le processus de normalisation postquantique.

2.5 Traitement et confidentialité des données dans l'IoT

L'IoT permet la collecte généralisée de données, le suivi de divers systèmes physiques et l'enregistrement, si besoin en temps réel, d'observations environnementales, de processus industriels ou d'autres activités humaines.

Ainsi, même si la collecte et l'exploitation de ces données permettent des avancées dans de nombreux domaines (comme la santé et le développement durable, pour n'en citer que quelques-uns), la protection de la vie privée devient un enjeu majeur, qui soulève des défis tant au niveau politique qu'aux niveaux scientifique et technique.

Sur les plans politique et juridique, de nouvelles réglementations peuvent permettre de résoudre en partie les problématiques liées à la vie privée. Parmi les exemples notables, on peut citer des cadres tels que le Règlement général sur la protection des données (RGPD) de l'Union européenne, les recommandations du WP29 et des directives analogues en matière d'éthique et de fiabilité.

Sur les plans scientifique et technique, des enjeux spécifiques deviennent prioritaires pour concevoir des techniques d'exploitation des données issues de l'IoT qui sont capables de protéger intrinsèquement la vie privée.

↗ L'équipe-projet **PRIVATICS** d'Inria étudie le suivi et la caractérisation de l'exposition des données personnelles dans des cas d'utilisation d'objets connectés, et conçoit des mécanismes qui améliorent la transparence pour les utilisateurs d'objets connectés et assurent leur consentement approprié.

Systèmes de protection de la vie privée

La collecte de données et la science des données (data science) sont au cœur des systèmes d'information modernes. Les objets connectés jouent aujourd'hui un rôle clé dans la collecte de ces données.

Cependant, l'approche traditionnelle - centralisée – pour la collecte des données, conduit à une impasse en matière de protection de la vie privée, ainsi qu'à une charge excessive du réseau lorsque la quantité de données à transférer est trop importante. De nouveaux paradigmes sont donc nécessaires, et les recherches actuelles explorent des alternatives.

Par exemple, au lieu de partager les données brutes, une entité peut partager des données prétraitées. Cette approche peut fonctionner dans différents modèles d'organisation, notamment des modèles client-serveur ou totalement décentralisés (pair à pair). La motivation des fournisseurs de données est de contrôler totalement la distribution des informations découlant de leurs données. Cela nécessite de répartir certains coûts de calcul, de stockage ou de communication entre un plus grand nombre de machines dans différentes parties du réseau, potentiellement à sa périphérie.

Pour permettre le prétraitement, il faut d'abord être en mesure de programmer les objets connectés. Il est donc nécessaire de disposer au préalable de plates-formes logicielles embarquées appropriées offrant une base suffisante, à la fois en termes d'ouverture et de performance. Les approches de partage de l'utilisation des données peuvent alors s'appuyer sur des techniques telles que :

- ***la transformation des données*** (par exemple, en ajoutant du bruit) pour obtenir certaines garanties d'anonymat (*differential privacy*, K-anonymat, L-sensibilité, etc.) ;
- ***des primitives cryptographiques spécifiques*** telles que le cryptage homomorphe pour protéger les données dans les calculs multipartites, etc ;
- ***la compression de données avec perte pour brouiller partiellement les informations,*** par exemple en ne communiquant que des agrégats de données, des modèles de prédiction (partiels), ou des calculs intermédiaires comme les gradients ou les statistiques, ou des combinaisons de ces éléments ;
- ***l'intégration de techniques de gestion des données,*** afin d'externaliser sélectivement les résultats (par exemple, émettre une alerte en fonction de la survenue d'une conjonction d'événements, plutôt que de l'ensemble des événements collectés) et d'agréger les informations collectées (par exemple, émettre une statistique calculée, plutôt que de vastes jeux de données brutes) ;
- ***des techniques de calcul sécurisé*** embarquées dans (et distribuées sur) des dispositifs matériels IoT sécurisés, pour garantir que le code de traitement attendu a bien été utilisé, avec les données d'entrée appropriées, afin de produire le résultat.

Le défi principal en matière de confidentialité consiste à concevoir des mécanismes qui maximisent l'utilité des données tout en protégeant la vie privée. Un autre défi en la matière est de minimiser le rapport coût/bénéfice des mécanismes de confidentialité pour les appareils fournisseurs de données (qui peuvent être des objets connectés avec des ressources très limitées).

↗ L'équipe-projet **PRIVATICS** d'Inria travaille à la conception et à la mise en œuvre d'algorithmes pour le partage d'informations protégeant la vie privée, qui sont applicables aux objets connectés portables allant des smartphones aux jetons intelligents.

↗ L'équipe-projet **ARIC** d'Inria travaille à la conception et à l'analyse d'algorithmes de cryptage entièrement homomorphes, et à leur utilisation pour des calculs protégeant la vie privée.

↗ L'équipe-projet **COMETE** d'Inria travaille à la conception et à l'analyse de mécanismes de confidentialité différentielle (differential privacy) utilisés pour des calculs protégeant la vie privée.

Smart container, solution pour le suivi et le monitoring des conteneurs. © *Photo Raphaël de Bengy.*

L'apprentissage automatique décentralisé avec l'IoT

Partager l'utilisation des données plutôt que les données elles-mêmes est un principe qui peut permettre de limiter les enjeux de protection de la vie privée, au niveau de l'apprentissage automatique (AA), qui exploite des données issues de l'IoT qui sont potentiellement confidentielles. Dans ce contexte, l'apprentissage fédéré (AF) fait l'objet de recherches actives[1] : son principe est qu'un certain nombre de clients collaborent via un serveur central pour entraîner un modèle, chacun conservant ses données d'apprentissage propres. Il s'agit là aussi de minimiser la collecte de données dans le but d'éliminer à la fois les problèmes de confidentialité et les goulots d'étranglement du réseau (lorsque la quantité de données à transférer est trop importante [2]). Un enjeu fondamental en la matière est l'absence d'autorité centrale contrôlant et distribuant les données entre les pairs. Comme les données restent sur la machine où elles ont été collectées, les algorithmes AA qui sont au cœur des systèmes de décision doivent traiter des données **non iid** (indépendantes et identiquement distribuées) ce qui est un challenge. En effet, la notion de « distribution identique » est l'hypothèse principale sur laquelle l'AA s'appuie pour garantir que le modèle entraîné se comportera comme attendu pour des valeurs futures, et la propriété « d'indépendance » simplifie considérablement la complexité des catégories de modèles possibles.

D'autres travaux de recherche explorent le principe de l'apprentissage fédéré entièrement décentralisé, c'est-à-dire sans serveur central, dans l'esprit des systèmes de type pair-à-pair. Dans ce contexte, il semble opportun de tirer parti du nombre important de pairs pour améliorer la confidentialité. Un enjeu clé consiste toutefois à définir avec qui collaborer et comment optimiser les coûts de communication. C'est là qu'apparaissent des problématiques de sécurité et de confiance typiques du pair-à-pair, qui doivent être revues et résolues dans ce contexte (par exemple en détectant les pairs malveillants et la collusion entre pairs). Des problèmes spécifiques se posent également selon les techniques de protection de la vie privée utilisées : par exemple, si l'on ajoute du « bruit » pour brouiller partiellement les données partagées, l'utilité peut diminuer radicalement lorsque les données disponibles sont trop peu nombreuses. Dernière remarque, et non des moindres : le calcul et l'état (du modèle et des données d'apprentissage) doivent être réduits au strict minimum afin de s'adapter aux faibles ressources disponibles sur les pairs qui s'avèrent être des appareils IoT bas de gamme..

1. Peter Kairouz et al. *Advances and Open Problems in Federated Learning*. Rapport technique, arXiv:1912.04977, https://arxiv.org/abs/1912.04977, 2019.
2. Par exemple, les voitures autonomes équipées de caméras et de capteurs collectent d'énormes quantités de données et, dans de nombreux endroits, la connexion au réseau est intermittente.

↗ L'équipe-projet **MAGNET** d'Inria travaille à la conception de méthodes de *machine learning* (apprentissage automatique) respectueuses de la vie privée qui exploitent des techniques d'anonymisation des données pour alimenter l'apprentissage, et des algorithmes pair-à-pair entièrement décentralisés, atténuant les hypothèses nécessaires pour permettre l'apprentissage fédéré.

Calcul décentralisé et bases de données sécurisés dans l'IoT

Le traitement des bases de données, et en particulier le big data, revêtent une importance capitale dans le contexte de l'IoT. Le partage de « l'utilisation » des données, plutôt que des données elles-mêmes, conduit à un nouveau paradigme, dans lequel les opérations de traitement des données sont délocalisées à la périphérie du réseau voire, dans les cas extrêmes, à bord des objets connectés eux-mêmes. Une telle approche favorise la confidentialité des données et le respect de la vie privée (notamment en appliquant localement des règles de contrôle d'accès et de confidentialité), ainsi que les économies d'énergie (en évitant la transmission de données rarement utilisées et en produisant des résultats agrégés au lieu de partager toutes les données brutes collectées).

Du point de vue des bases de données, elle amène à considérer que les vastes ensembles d'objets connectés (dotés de ressources de stockage et de calcul) constituent une base de données distribuée ou fédérée, sur laquelle on peut lancer un traitement global. Cette vision met les chercheurs face aux défis suivants :

- (1) rendre les techniques liées aux bases de données (stockage et indexation, algorithmes d'évaluation des requêtes) compatibles avec les contraintes matérielles sévères imposées par les objets intelligents, en particulier, la mémoire vive très limitée face à une mémoire flash relativement importante ;
- (2) concevoir, pour les objets connectés, de nouvelles techniques sécurisées d'évaluation des requêtes distribuées, disponibles à une très grande échelle sans recourir à un serveur central fiable.

Le premier défi découle des contraintes matérielles conflictuelles des objets connectés en matière de techniques de gestion des données : la mémoire vive limitée nécessite l'indexation massive des données (car il n'est pas possible de construire des résultats intermédiaires dans la mémoire vive pendant l'exécution), mais la spécificité de la mémoire flash NAND pénalise les petites écritures aléatoires (qui sont nécessaires pour maintenir les index). Des travaux pionniers sont menés

chez Inria pour lever ces contraintes en permettant, grâce à de nouveaux principes de conception, de stocker et de traiter dans un objet intelligent des millions d'entrées de base de données enregistrées. Le prochain défi pour les chercheurs sera de généraliser ces résultats aux séries de données, qui constituent le principal type de données présent dans les objets connectés.

Pour ce qui est du second défi, on peut établir un parallèle avec les fédérations de données privées, qui font l'objet de recherches actives en lien avec les bases de données, et dans lesquelles l'objectif est, pour un ensemble de propriétaires de données, de se servir de leurs propres données pour répondre à une requête globale, sans divulguer ces données (potentiellement sensibles) aux autres. Plusieurs approches faisant appel à différentes techniques comme la cryptographie (calcul multipartite sécurisé), l'ajout de bruit (confidentialité différentielle) ou l'informatique de confiance (sécurité matérielle) sont actuellement à l'étude. La transposition de ces techniques dans le contexte de l'IoT est difficile, car elle nécessite de prendre en compte d'immenses ensembles de « propriétaires de données » (des millions d'objets intelligents sont potentiellement concernés) avec des contraintes de ressources et d'énergie. Certaines études préliminaires s'appuient sur les composants matériels sécurisés présents dans certains objets connectés (comme les puces sécurisées ou les TPM) pour traiter des calculs de type *big data* (par exemple *MapReduce*) avec des garanties de confidentialité et d'intégrité, mais le sujet s'inscrit dans une perspective de recherche plus large.

Plus généralement, les solutions apportées à ces défis permettent de redonner aux individus la maîtrise de leurs données personnelles, et d'aider les citoyens à contribuer collectivement à des processus de traitement des bases de données de toute nature (SQL, *Big Data*, IA, etc.), et ce, de manière sécurisée, sans nécessairement recourir à un tiers de confiance.

↗ L'équipe-projet **PETRUS** d'Inria travaille à la conception d'une architecture de base de données fiable pour les systèmes de gestion de données personnelles (SGDP, avec des garanties de sécurité et de confidentialité, pour permettre des requêtes distribuées impliquant de très grands ensembles de SGDP. **PETRUS** a notamment développé PlugDB, une plate-forme matérielle et logicielle de gestion de données personnelles, qui utilise des technologies de calcul sécurisé et des microcontrôleurs.

2.6 Sûreté, fiabilité et certification pour l'IoT

Une fois pleinement déployé, l'IoT va accroître notre dépendance vitale à l'égard des systèmes embarqués, ainsi que des capteurs et des actionneurs mis en réseau. Dans de nombreuses applications, le terme « vital » est utilisé littéralement, car l'IoT peut déclencher des effets physiques directs..

Une génération d'implants intelligents est ainsi en train d'émerger, impliquant un équipement embarqué, une boucle locale sans fil et un code source ouvert. Le système de pancréas artificiel est par exemple conçu pour adapter automatiquement l'activation d'une pompe à insuline afin de maintenir la glycémie à un taux approprié pendant la nuit et entre les repas. Et tout dysfonctionnement de ce système met immédiatement en péril la vie du patient.

De façon plus générale, une série d'incidents liés à la sécurité et à la fiabilité des objets connectés ont récemment impacté un certain nombre de machines, notamment des véhicules comme des voitures connectées et même des avions. Du fait de défaillances de ces appareils connectés, des utilisateurs ont vu leur vie gravement menacée, ou l'ont perdue à la suite d'accidents réels. Considérée collectivement, la série d'incidents récents laisse à penser que :

- même un code axé sur une sécurité extrême peut contenir des bugs fatals et/ou exploitables ;
- même un code-source fermé et hautement sensible peut faire l'objet d'une fuite.

Alors que les utilisations des objets connectés s'étendent au-delà de simples « gadgets amusants », les aspects liés à leur sûreté prennent une importance fondamentale. L'utilisation potentielle des objets connectés de façon inattendue ou dans des contextes potentiellement hostiles et propices aux cyberattaques pose certains problèmes. Satisfaire simultanément aux exigences de sécurité et de sûreté dans le domaine de l'IoT constitue un défi de taille. La certification en est un exemple parlant : bien souvent, l'objet connecté est une cible mouvante (du fait des mises à jour logicielles) et le contexte dans lequel il est utilisé n'est pas restrictif (notamment dans le cas de l'électronique grand public). Dès lors, que faut-il certifier exactement, et comment ?

Pour relever les défis dans ce domaine, les communautés de chercheurs sur la sûreté et la sécurité - qui sont traditionnellement des communautés plutôt distinctes – doivent collaborer plus étroitement, et repenser ensemble les **concepts fondamentaux** depuis le niveau le plus basique.

D'une part, garantir la sûreté et la sécurité n'est plus une tâche « individuelle », mais implique un ensemble complexe de parties prenantes : fournisseurs de logiciels, opérateurs, régulateurs, utilisateurs individuels, etc. Dans un tel contexte, les cadres juridiques doivent être en mesure de déterminer qui est responsable de quoi.

D'autre part, l'interdépendance accrue entre les systèmes oblige à reconsidérer ce que nous entendons par « critique ». Par exemple, des études ont montré qu'un *botnet* de taille moyenne composé de climatiseurs et de radiateurs (activables à distance en tant qu'objets connectés) peut être utilisé comme une arme pour perturber le réseau électrique national. La sûreté et la sécurité des systèmes mixtes, comprenant non seulement des composants critiques classiques, mais aussi des composants connectés grand public, constituent donc un défi. La certification nécessaire de ces composants est une tâche complexe, non seulement parce qu'elle est difficile à formaliser dans ce contexte (qu'est-ce qu'une cybersécurité suffisante ?), mais aussi parce qu'elle doit rester très rentable, puisqu'il s'agit de dispositifs à faible coût.

Cartographie de la salle IoT du Lab de Lyon. © *Photo C. Morel.*

2.7 Interaction Homme-machine avec l'IoT

Avec l'IoT, les ordinateurs « disparaissent » de plus en plus, et une interaction de machine à machine (M2M) toujours plus complexe s'établit en coulisse. Grâce aux capteurs et aux actionneurs, et alors que l'interaction de l'humain avec le système prend de nouvelles formes (elle peut, par exemple, être basée sur des gestes), la réalité physique elle-même peut être personnalisée et vécue différemment. La postphénoménologie étudie l'interaction entre les humains, le monde et les technologies modernes, ces dernières étant des médiateurs non neutres. Paradoxalement, avec le M2M, le facteur humain devient encore plus fondamental. Alors même que les humains sont court-circuités de façon plus systématique afin d'optimiser et d'automatiser, ils peuvent être davantage affectés par un dysfonctionnement du système ou par un manque de compréhension de son fonctionnement. Les ouvriers, par exemple, risquent de plus en plus d'être déqualifiés, voire remplacés.

Alors que l'IoT se généralise, il pourrait engendrer une nouvelle fracture numérique, dans laquelle certains utilisateurs sont fortement désavantagés par la technologie, alors que d'autres obtiennent une capacité de contrôle plus étendue, et que d'autres encore gagnent du pouvoir grâce au contrôle habile que permettent certaines API *(Application Programmine Interface)*. Dans ce contexte, il est essentiel de concevoir avec soin un nouveau type d'interaction Homme-machine.

Donner aux humains les niveaux de contrôle adéquats sur l'IoT (quel est le bon niveau ?)

Prenons l'exemple des thermostats. Les thermostats ont initialement été conçus pour commander la température, généralement avec des interfaces minimalistes comme un bouton rotatif. Ils ne font rien de plus. Avec l'IoT, ces appareils peuvent disposer de multiples autres fonctionnalités : ils peuvent être programmés, utilisés pour commander d'autres systèmes, reliés à d'autres appareils à des capteurs, ou à un service en ligne, etc. Cependant, bien que leur fonctionnalités soient enrichies de la sorte, beaucoup d'objets connectés conservent leur interface minimaliste d'origine. Cette approche (appelée *retrofitting*) soulève un certain nombre de questions. D'une part, ces interfaces sont simples et familières. D'autre part, les interfaces de substitution de base (par exemple, l'écran d'un *smartphone*) ne permettent pas d'exploiter le plein potentiel des interactions cyberphysiques. Bien

que les aspects technologiques de l'intégration de l'IoT dans le monde réel soient bien avancés, il nous reste à concevoir une transition fluide entre ces deux mondes.

Comprendre ce qu'il se passe « derrière » l'IoT

Pour atteindre un niveau de transparence adéquat, les utilisateurs doivent avoir une certaine compréhension de l'IoT. Les infrastructures IoT sont complexes par nature : elles sont composées de machines hétérogènes interconnectées et de dispositifs interactifs, sous-tendus par une « intelligence ambiante ». Cette complexité soulève des défis en termes d'interaction, pour que les utilisateurs puissent contrôler ces systèmes. En matière de technologies interactives, la tendance générale de ces dernières décennies a été de simplifier à l'extrême les interfaces pour faciliter l'interaction. Elle a certainement contribué à démocratiser l'utilisation de la technologie pour les néophytes, mais au prix d'une moindre expressivité. Dans le contexte de l'IoT, cette approche ne pourra probablement pas s'intensifier en raison de la complexité des infrastructures et des moyens d'interaction particuliers qui sont offerts (par exemple appareils multiples,entrées/sorties réduites, interaction distante et distribuée).

L'un des enjeux consiste donc à mieux prendre en compte la complexité en concevant des interfaces et des interactions adaptées, afin que les utilisateurs puissent progressivement élaborer un modèle mental approprié pour le système, ce qui implique :

- comprendre et anticiper la façon dont le système va réagir à leurs actions ;
- avoir une vision claire et correcte des états du systèmes et des erreurs pouvant survenir ;
- acquérir progressivement des compétences pour contrôler finement le système.

Le contrôle partagé constitue un autre enjeu. Les systèmes IoT disposent souvent d'un certain degré d'autonomie et peuvent prendre l'initiative d'effectuer des tâches ou de proposer des actions aux utilisateurs. Cet aspect peut devenir critique. Les utilisateurs doivent connaître l'état présent du système, car il peut avoir évolué de manière autonome ; ils doivent savoir comment reprendre le contrôle en cas de besoin ; et quand le système agit de lui-même, ils doivent avoir le sentiment de maîtriser les choses. Les utilisateurs doivent en particulier pouvoir :

- compter sur une cohérence du système, confronté à des situations similaires ;
- identifier les actions appropriées et pouvoir communiquer leurs intentions ;
- surveiller le système de manière que les erreurs puissent être identifiées et corrigées.

Les crashs du Boeing 737MAX, qui étaient en partie imputables à la dissimulation d'un comportement non-documenté du système, sont des exemples de contrôle partagé qui ont horriblement mal tourné. Des travaux de recherche dans le domaine de l'interface Homme-machine (IHM) ont commencé à analyser des incidents de ce type, afin de caractériser la compréhension et la visibilité des états du système au niveau des utilisateurs.

L'un des pièges en la matière est une situation dans laquelle le système contrôle l'utilisateur davantage que l'utilisateur ne contrôle le système. Une approche basée sur l'intelligence computationnelle pourrait être utilisée pour anticiper les actions de l'utilisateur, par le biais du suivi et de l'inférence des tâches, supprimant ainsi la nécessité d'avoir une interface explicite et activée en permanence. La technologie ubiquitaire, explorée pour la première fois dans les travaux de recherche de Wellner avec son Bureau Numérique, promettait notamment des environnements qui anticipent l'utilisateur. En bref, grâce à des modèles d'utilisateurs et de tâches, l'environnement peut se reconfigurer de manière autonome, tout en permettant à l'utilisateur de garder le contrôle. Enfin, cette catégorie d'approche pourrait également permettre d'adapter les interfaces à différents types de publics, en fournissant aux utilisateurs des niveaux de contrôle appropriés en fonction de leurs besoins, de leurs compétences et de leurs contextes d'utilisation.

↗ Des équipes-projets d'Inria dont **AVIZ**, **EXSITU**, **ILDA** et **LOKI** étudient et conçoivent de nouvelles méthodes d'interaction et de nouveaux systèmes interactifs qui renforcent l'autonomie des utilisateurs en prenant davantage en compte leurs capacités et de leur expertise.

Exploiter le potentiel de l'interaction cyberphysique

Les travaux de recherche sur l'interaction tactile et tangible révèlent un potentiel considérable pour étendre le champ d'interaction Homme-machine (jusqu'à présent assez réduit) *via* les objets connectés. En effet, la recherche dans le domaine de l'interaction Homme-machine ne cesse d'explorer des techniques de détection avancées capables d'identifier et de différencier nos façons d'entrer en contact avec une surface (par exemple l'identification des doigts, la détection précise du point de contact, la pression appliquée, l'inclinaison du doigt). De même, la manière de saisir et de manipuler un objet physique peut donner beaucoup d'indications sur l'intention de l'utilisateur. Imaginons, par exemple, une trousse à crayons : on la saisira différemment selon que notre intention est de l'ouvrir, de la ranger ou de la donner à quelqu'un d'autre. Ce principe s'applique à d'autres

objets, et les appareils connectés ne font pas exception. Associées à des études sur la capacité des utilisateurs à tirer profit des technologies de détection tactiles et tangibles, ces recherches visent à améliorer la portée de l'interaction entre les utilisateurs et les objets connectés, sans interface additionnelle/externe : un seul bouton pourrait par exemple déclencher des actions différentes selon la manière dont on le touche ou le saisit.

Cependant, ces approches ne répondent que partiellement aux problématiques de visibilité limitée concernant les différentes actions possibles avec le système, et les manquements du système en termes de *feedforward* (ce qu'il faut faire) ou de *feedback* (ce qui a été fait). D'une part, la possibilité de rendre l'interaction plus « physique » pourrait aider l'utilisateur à transférer des connaissances issues d'autres contextes. D'autre part, l'augmentation des capacités d'interaction (par exemple en ajoutant la détection tactile à un bouton physique) pourrait contribuer à améliorer la visibilité et l'accessibilité des actions et des fonctionnalités disponibles. Sur ce plan, la réalité mixte (RM) est un moyen prometteur pour rendre plus apparentes ces « interfaces invisibles », par exemple avec des smartphones ou des dispositifs portables tels que des lunettes, en incrustant des signes discrets d'interaction sur les dispositifs physiques, voire en superposant des didacticiels complets, à la demande de l'utilisateur.

La RM pourrait également offrir aux utilisateurs un meilleur *feedback* de la part du système, susceptible aussi d'être complété par des solutions haptiques comme le *feedback* vibrotactile. Des travaux de recherche très actifs dans le domaine de l'interaction Homme-machine utilisant l'interaction avec les objets connectés visent à trouver de nouveaux moyens pour réaliser ce type de *feedback* sur n'importe quelle surface (par exemple des actionneurs, l'électrovibration, etc.). On cherche à comprendre comment ce *feedback* est perçu par les utilisateurs et quels types et quantités d'informations il peut transmettre. Un domaine connexe concerne la détection du mouvement humain pour supporter un large éventail d'applications de réadaptation et de création.

↗ Des équipes-projets d'Inria dont **AVIZ**, **EXSITU**, **ILDA** et **LOKI** explorent de nouveaux matériaux et dispositifs interactifs afin de créer des formes inédites d'interfaces tangibles pour une grande variété d'applications domestiques, professionnelles et créatives.

2.8 Contrôle avec l'IoT dans la boucle

Un objectif majeur de l'IoT est de permettre la supervision et le contrôle avancés de systèmes distribués déployés dans divers environnements : des maisons et immeubles connectés, aux villes intelligentes, en passant par l'industrie 4.0. Par le biais du contrôle optimisé qu'il permet, l'IoT peut notamment offrir des capacités de supervision et des performances optimales aux systèmes de petite taille qui ne peuvent pas se permettre d'employer de dispositif de contrôle dédié.

L'utilisation de réseaux partagés polyvalents pour contrôler des éléments répartis géographiquement conduit à des architectures très flexibles. L'inconvénient est que des dynamiques asynchrones sont ajoutées dans les boucles, ce qui peut fortement dégrader les performances, voire la stabilité. L'évaluation des effets du réseau et la conception de systèmes de contrôle en réseau (*networked control systems*, ou NCS) robustes constituent un défi, qui suppose de modéliser ces systèmes hybrides en associant des états continus et des événements discrets, des fonctions de temporisation, etc. Les algorithmes doivent également être capables de gérer des hiérarchies complexes de sous-systèmes impliquant de telles dynamiques asynchrones, de multiples niveaux, des échelles de temps extrêmement diverses (du mois à la microseconde) et une portée géographique très variée (de « systèmes sur une puce » à des dispositifs de taille planétaire). Concevoir une commande en boucle fermée sur de tels réseaux non déterministes requiert une résilience extrême face aux variations (inévitables) en termes de latence, de gigue et de débit.

↗ L'équipe-projet **VALSE** d'Inria travaille à la modélisation et à l'analyse de systèmes hautement dynamiques, distribués et incertains, que l'on rencontre dans l'IoT et les systèmes cyberphysiques. **VALSE** conçoit des algorithmes d'estimation et de contrôle décentralisé robustes en utilisant les concepts de temps fini/ temps fixé/ convergence hyperexponentielle.

Dans ce domaine, un enjeu connexe est la **gestion autonome des boucles de rétroaction dans les intergiciels IoT.** Dans l'IoT, le contrôle et la surveillance requièrent généralement l'utilisation d'intergiciels pour la supervision et la gestion de l'infrastructure. Les intergiciels IoT doivent permettre la gestion (centralisée ou décentralisée) de composants logiques complexes et distribués, dans des infrastructures très hétérogènes : par exemple, des appareils de petite taille avec une puissance de calcul limitée, des passerelles domotiques, des nœuds locaux

du réseau cellulaire, ou des centres de données dans le *Cloud*. Les intergiciels IoT doivent donc fournir des abstractions utilisables pour une très grande diversité de systèmes d'exploitation et de protocoles de communication.

L'automatisation des boucles de rétroaction cyberphysiques, par exemple avec l'informatique autonome, est un défi fondamental sur ce plan. Ces boucles de contrôle doivent permettre une auto-adaptation continue du système cyber-physique, qui va réagir aux informations observées en prenant des décisions sur la base d'une représentation du système, mises en œuvre au travers d'actions appliquant une stratégie ou une politique définies au préalable à haut niveau. Le but est de contrôler des variations de dynamique (potentiellement importantes) qui se produisent soit dans l'environnement physique supervisé (qui est l'objet habituel de la théorie du contrôle), soit directement dans l'infrastructure du système de calcul et de communication (par exemple des variations de charge, la tolérance aux pannes, l'autoprotection).

↗ L'équipe-projet **ACENTAURI** d'Inria étudie de nouveaux paradigmes qui augmentent l'autonomie des systèmes robotiques, permettent un comportement axé sur les tâches, et exploitent la perception et le contrôle multisensoriels.

Les enjeux dans ce domaine comprennent aussi la conception et l'optimisation des mécanismes de reconfiguration automatique et des architectures logicielles pour les aspects fonctionnels liés aux applications ainsi que pour les aspects calculs liés à l'infrastructure (par exemple migration des services, auto-adaptation pour passage à l'échelle). Entre ces différents niveaux, des exigences de séparation des problématiques existent également.

↗ L'équipe-projet **CTRL-A** d'Inria travaille à la conception de méthodes pour du contrôle utilisé dans le cadre de l'informatique autonome, en s'appuyant sur la théorie du contrôle, pour permettre une gestion adaptée aux applications des architectures informatiques reconfigurables, dans le domaine de l'IoT et dans celui du calcul haute performance (HPC).

Combler le fossé entre la robotique et l'IoT pour l'industrie

Les essaims coordonnés de minuscules robots (microrobots) présentent un intérêt particulier permettant de nouvelles applications en robotique. Ils ont le potentiel pour surpasser les robots monolithiques dans les applicationspour

lesquelles la diversité spatiale présente des avantages, comme la détection distribuée. La microrobotique en essaim est une nouvelle frontière pour la recherche sur l'IoT industriel, car elle nécessite de répondre simultanément à un certain nombre de problèmes ouverts, afin de permettre le contrôle et l'interaction impliquant un grand nombre de microrobots.

Un premier enjeu est de permettre la mobilité. Si les protocoles de communication IoT à basse consommation énergétique sont désormais largement normalisés et en cours de déploiement, ils ont surtout été conçus pour interconnecter des dispositifs relativement statiques dans une zone donnée. Le déplacement d'une partie ou de la totalité de ces dispositifs n'est pas pris en charge efficacement par les protocoles IoT industriels standard actuels (par exemple *6TiSCH*).

Plateforme expérimentale FIT (Future Internet of Things). © Inria / Photo C. Morel.

Un deuxième enjeu est d'avoir une latence faible et prévisible. Les réseaux IoT industriels sont actuellement en mesure de garantir la transmission des données générées, mais seulement à destination d'une passerelle, par exemple. Si la latence n'est jamais un paramètre bien défini (le sans-fil étant aléatoire), la planification prévue dans le réseau peut la rendre prévisible. Ce déterminisme est nécessaire pour l'exécution de boucles de contrôle via les réseaux IoT.

Un troisième enjeu est d'assurer une localisation précise et parcimonieuse. Les fondements pour des techniques de localisation telles que les méthodes

d'évaluation de l'angle d'arrivée UWB ou BLE ont été mis au point. Un sujet de recherche dans ce domaine consiste à coconcevoir la solution de localisation et les protocoles de communication que permettra une localisation à la demande, compatible avec les exigences de faible consommation énergétique de la plupart des applications IoT.

↗ L'équipe-projet **EVA** d'Inria travaille sur la robotique en essaim expérimentale. **EVA** a par exemple mis au point Atlas, un simulateur de robotique en essaim, et conçoit actuellement *DotBot*, un grand banc d'essai pour des essaims de robots à l'échelle centimétrique, pouvant compter jusqu'à mille unités. L'équipe-projet **AVIZ** travaille à la conception d'essaims de minuscules robots capables d'effectuer des visualisations physiques et une variété d'autres tâches. Elle a notamment mis au point la plate-forme *Zooids*.

2.9 La sécurité dans l'IoT

La plupart des crimes commis actuellement impliquent des composants cyberphysiques. D'autre part, les cyberattaques impliquant des entités au-delà des frontières nationales constituent désormais notre réalité de tous les jours. Le piratage en ligne, qu'il soit motivé par le profit ou mené par les États, atteint des niveaux sans précédent : la « troisième guerre mondiale » a lieu en ligne.

Dans ce contexte, les enjeux de sécurité sont présents à tous les niveaux de l'IoT. La sécurité et la résilience d'un système dépendent de son maillon le plus faible : la sécurisation des objets connectés de faible puissance est donc cruciale.

Au-delà des fraudes de cybersécurité les plus basiques (*phishing*, ingénierie sociale, etc.), des attaques toujours plus variées nécessitent la mise en œuvre de mesures spécifiques pour atténuer les risques et de nouveaux mécanismes de sécurité à tous les niveaux du système.

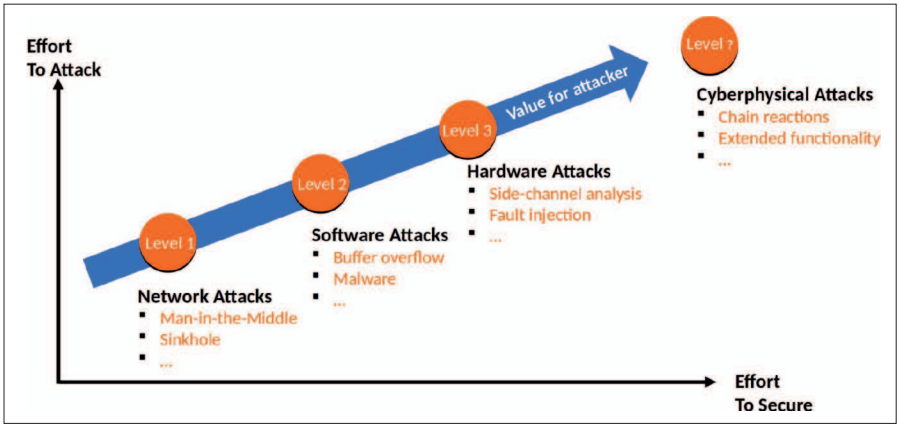


Figure 1 : Surface d'attaque dans l'IoT.

Modélisation des attaquants dans l'IoT

Les modes opératoires traditionnels des cyberattaques restent efficaces dans l'IoT : série de tentatives pour exploiter différentes vulnérabilités, augmentation progressive de privilèges usurpés, etc. Toutefois, *l'évaluation des risques est considérablement différente avec l'IoT.* Si de plus en plus d'actionneurs contrôlés *via* l'IoT vous entourent et influencent physiquement votre environnement (voire votre état biologique, comme les implants intelligents), le niveau de risque que

vous tolérez est d'autant plus faible. Si des capteurs améliorés par l'IoT collectent des données de plus en plus intimes et fines (comme votre rythme cardiaque et votre taux de transpiration détectés en temps réel), l'impact des atteintes à la vie privée est d'autant plus important.

La sécurité informatique classique doit déjà prendre en compte une surface d'attaque énorme, qui va des attaques réseau (*man-in-the-middle, sinkhole,* etc.) aux attaques logicielles (logiciel malveillant, dépassement de mémoire tampon) et matérielles (injection de fautes, attaques par canaux auxiliaires, etc.), en passant par le vecteur humain et les méthodes d'ingénierie sociale. *Mais avec l'IoT, de nouveaux vecteurs d'attaque apparaissent.*

Il est désormais possible de déclencher à distance des réactions en chaîne cyberphysiques à conséquences catastrophiques, des effets dominos qui exploitent massivement les botnets et l'interdépendance accrue entre des systèmes autrefois isolés les uns des autres, comme le réseau électrique et Internet par exemple. D'autre part, des attaques par extension de fonctionnalité permettent d'armer un dispositif hostile contrôlé *via* l'IoT en détournant son utilisation d'une manière totalement inattendue. Ces attaques élargissent la surface d'exposition des systèmes distribués.

En outre, l'apparition de nouvelles interfaces utilisateur cyberphysiques offre également de nouveaux vecteurs d'attaque. Par exemple, l'émergence des commandes vocales permet des attaques nouvelles contre les assistants personnels vocaux, l'authentification de l'assistance vocale par l'utilisateur (ou vice-versa) étant difficile et pouvant facilement faire l'objet d'abus. Dans un avenir proche, on s'attend à ce que les nouvelles interfaces utilisateurs cyberphysiques intègrent des implants intelligents sophistiqués semblables au prototype d'interface cyberphysique cérébrale *Neuralink*, ce qui augmentera encore les enjeux en termes d'exigences de sécurité et de sûreté.

La définition de nouveaux modèles d'attaquants prenant en compte ce contexte est donc un défi crucial.

Sécurisation des protocoles réseaux dans l'IoT

Les avantages apportés par l'IoT reposent sur l'intégration via le réseau de nouveaux appareils qui étaient auparavant absents, ou qui fonctionnaient de manière autonome et isolée. Le revers de la médaille, c'est qu'ils ouvrent de nouvelles voies d'accès aux cyberattaques par le réseau. Il est donc fondamental de sécuriser les communications réseau dans l'IoT.

La pile de communication réseau est traditionnellement divisée en couches d'abstractions imbriquées. Chaque couche fournit des services à la couche supérieure, et utilise les services de la couche située directement en-dessous. Le modèle dominant est celui de l'actuel Internet, composé des couches application, transport, réseau, liaison de données et physique. Des mécanismes de sécurité spécifiques sont nécessaires au niveau de chacune des couches de la pile de protocoles réseau.

Pour bien appréhender les défis à relever dans ce contexte, il faut d'abord rappeler certaines particularités des dispositifs et des réseaux de l'IoT, en comparaison avec des machines ordinaires connectées à la périphérie du réseau Internet.

DÉBIT DE TRANSFERT DE DONNÉES TRÈS FAIBLE

La mise en réseau locale dans l'IoT (ou boucle locale IoT) repose souvent sur des ondes radio basse puissance, qui présentent des contraintes inhabituelles en termes de débits de données physiques : de 250 kilobits par seconde annoncés pour les technologies à courte portée (connectivité de l'ordre des dizaines de mètres en intérieur et de centaines de mètres en extérieur), à 10 kilobits par seconde environ pour les technologies longue portée (connectivité allant jusqu'à 10 km en extérieur). Par comparaison, ces valeurs représentent *grosso modo* entre 0,01 % et 0,1 % des débits de données annoncés pour le WiFi moderne ou la 4G cellulaire.

PATTERNS SPÉCIFIQUES POUR LE TRANSFERT DE DONNÉES

Les données générées par les objets connectés sont souvent stockées à des niveaux intermédiaires avant d'arriver au consommateur final. Il ne suffit donc plus de faire confiance aux données en se basant sur l'identité d'un pair communicant, comme c'est souvent le cas dans l'Internet traditionnel. Au lieu de cela, un modèle producteur-consommateur est nécessaire à des fins de sécurité, pour offrir des garanties de sécurité au niveau de la couche application.

BUDGETS MICROSCOPIQUES POUR LES RESSOURCES EMBARQUÉES

Les objets connectés ont des budgets très faibles en termes de puissance, de traitement ou de mémoire. La mémoire, par exemple, représente 0,001 % du budget ressources disponible sur une machine classique connectée à Internet.

FACTEUR HUMAIN DIFFÉRENT

Les objets connectés de faible puissance sont souvent dépourvus d'interfaces utilisateur courantes telles qu'un écran ou un clavier. Les boutons-poussoirs et les LED étant les seuls moyens d'interagir avec un appareil, l'étape de configuration sur le terrain (à des fins de mise en service ou de débuggage) devient nettement

plus difficile. En outre, les objets connectés ont généralement un rapport Homme-objet intrinsèque de 1:N (en considérant tous les capteurs/actionneurs/implants et gadgets intelligents), alors que d'autres machines connectées à la périphérie du réseau tendent plutôt vers un ratio de 1:1 (*smartphone*, ordinateur portable).

Amorcer la sécurité, sans interface utilisateur

Dans le cadre des solutions de sécurité des communications IoT basse consommation telles que définies par les organismes de normalisation, on suppose habituellement que la relation de confiance (sous forme d'une clé cryptographique commune) entre les différentes entités impliquées dans la communication a été préétablie. Au moment de la fabrication par exemple, une telle relation de confiance s'établit généralement entre l'objet connecté et son fabricant. Toutefois, le domaine dans lequel cet objet connecté va être installé n'est pas connu au moment de la fabrication. Avant que l'objet puisse rejoindre un domaine donné, il doit être doté d'informations d'identification supplémentaires et spécifiques. L'établissement de cette relation de confiance entre l'objet connecté et le propriétaire du domaine a jusque récemment été considéré comme hors du champ de compétence des organismes de normalisation. Il s'agit pourtant d'une tâche importante étant donné que la plupart des objets connectés ne disposent pas d'interfaces utilisateurs communes (écran, clavier, etc.). Demander un mot de passe à un objet connecté bas de gamme n'est tout simplement pas envisageable, et le recours à des mécanismes d'authentification automatisés est dès lors à privilégier. Les entreprises utilisent généralement des canaux hors-bande (par exemple, communication NFC, réseau sans fil *ad hoc*, clés prépartagées imprimées à l'arrière d'un appareil, port série etc.). Premièrement, cette approche ouvre la voie à diverses vulnérabilités car le protocole « d'amorçage » est souvent conçu en interne, sans examen approfondi de la communauté et des experts en sécurité. Deuxièmement, elle passe difficilement à l'échelle (imaginez devoir amorcer une communication sécurisée pour des dizaines de capteurs à la fois…). L'un des enjeux consiste donc à définir des protocoles d'amorçage appropriés en tenant compte, d'une part, des contraintes liées aux appareils et aux réseaux IoT, d'autre part des contraintes opérationnelles et du cycle de vie des objets connectés.

↗ L'équipe-projet **EVA** d'Inria travaille à la conception de protocoles de sécurité sans contact et à des évaluations de performance des candidats aux normes de sécurité des communications dans des cas d'utilisation spécifiques à l'IoT.

Sécurisation du paradigme réseau IoT axé sur les données

Les communications internet ont été conçues pour interconnecter des terminaux distants qui réagissent simultanément, et sécuriser des canaux de communication relativement durables, transportant des flux de données entre ces terminaux. L'approche traditionnelle (tant dans la communauté des normalisateurs que dans celle des chercheurs) a consisté à réduire le coût des communications par l'intermédiaire d'un codage plus efficace, sans toutefois compromettre le niveau de sécurité. Les chercheurs ont par exemple proposé des versions allégées des protocoles IPsec et (D)TLS à la couche transport, réduisant le coût des communications en comprimant les en-têtes de contrôle (non critiques pour la sécurité) de tels protocoles.

Dans une large mesure, cependant, la communication IoT induit un trafic différent, *dit orienté vers les données :* il s'agit par exemple d'une communication ponctuelle (comme une mesure périodique par capteur ou une mise à jour de logiciel) qui fait intervenir des machines qui peuvent être, la plupart du temps, en mode économie d'énergie (veille), de telle sorte que, quelque part au cours de leur transfert sur le réseau, les données IoT seront stockées et séjourneront temporairement dans un répertoire sur le chemin de leur destination.

Les exigences imposées par ce paradigme orienté vers les données n'ont été abordées que récemment dans le cadre de travaux visant à définir de nouveaux mécanismes basés sur les « primitives de sécurité » des objets connectés, qui appliquent les mécanismes de protection au niveau de la couche application. De nouveaux protocoles et mécanismes légers doivent être définis et normalisés. On peut citer à titre d'exemple les travaux de recherche menés dans le domaine de la conception d'un protocole réseau centré sur l'information (ICN) et de ses extensions de sécurité. Un autre exemple est l'activité en cours autour de la normalisation d'*OSCORE* pour la sécurité des objets, c'est-à-dire la protection du transfert web utilisant CoAP, et le protocole *EDHOC* pour l'échange de clés au niveau de la couche application. Ces catégories de solutions promettent un faible (et très attractif) coût des communications et une prise en charge native de la sécurité des données IoT lorsqu'elles « reposent » quelque part au cours de leur transfert via le réseau, ce qui constitue une amélioration majeure par rapport aux solutions traditionnelles.

Il s'avère néanmoins souvent délicat d'établir une base commune raisonnable pour comparer des protocoles. La capacité à réaliser des comparaisons pertinentes entre nouveaux protocoles IoT et protocoles traditionnels reste un défi permanent pour la communauté de recherche.

Sécurisation des spécifications de nouveaux protocoles IoT

De nombreuses solutions IoT (qu'il s'agisse des spécifications ou de leurs implémentations) ne sont apparues que récemment. La nouveauté relative de ces protocoles implique que, comparés à des protocoles éprouvés (comme TLS, par exemple), ils ont été soumis à un nombre d'analyses plus limité, non seulement au niveau des garanties de sécurité des spécifications contre les vulnérabilités, mais aussi sur le plan de la robustesse et de l'efficacité de l'implémentation, des optimisations algorithmiques etc.

Un enjeu majeur pour les chercheurs est donc l'analyse formelle de ces nouvelles solutions légères définies par les organismes de normalisation tels que l'IETF, pour prouver les garanties de sécurité annoncées, en lien avec les performances visées en termes d'efficacité énergétique.

↗ L'équipe-projet **PROSECCO** d'Inria travaille à la conception de méthodes de vérification formelle applicables à la spécification de protocole IoT à basse consommation énergétique.

Au final, c'est le rôle de l'administrateur réseau de décider de la combinaison de protocoles à utiliser, des implémentations à exécuter et de la façon dont elles doivent être configurées afin de limiter les menaces liées à un déploiement donné. Les compétences et les connaissances de base des administrateurs réseau sont variables, en particulier avec les protocoles plus récents. L'expérience montre que, souvent, les paramètres par défaut ne sont pas modifiés, ce qui conduit à des vulnérabilités bien connues au sein du système. Pour être exploitable, la sécurité dans l'IoT exige ainsi non seulement des spécifications solides, mais aussi la capacité à réaliser des implémentations donnant des garanties de **sécurité par défaut** adéquates sur **tous** les appareils. Le défi consiste donc à faire évoluer conjointement les spécifications des protocoles IoT et leurs mises en œuvre, afin d'obtenir la sécurité adéquate en pratique.

Détection et traitement des incidents de sécurité dans l'IoT

La sécurisation d'un système distribué passe par un autre aspect essentiel : la capacité à gérer le réseau et les services. Des audits globaux du système peuvent révéler des vulnérabilités potentielles avant qu'elles ne soient exploitées. La surveillance du système permet de détecter les incidents de sécurité lorsqu'ils surviennent. Mais lorsque le système comporte des composants IoT cyberphysiques, sa complexité et son hétérogénéité croissantes font naître des enjeux particuliers.

L'un de ces enjeux réside dans la collecte (balayage passif ou actif) d'informations pertinentes permettant de suivre efficacement les caractéristiques et l'activité des objets connectés « *in vivo* », ce qui est particulièrement difficile à réaliser dans les limites des ressources de ces objets. Sur ce plan, de nouveaux protocoles doivent être développés et/ou instrumentés, afin de fournir des points d'observation appropriés dans les déploiements IoT.

Un autre enjeu consiste à automatiser la vérification croisée des informations d'audit spécifiques collectées dans un déploiement donné, par le biais de bases d'informations de sécurité maintenues au niveau mondial (CPE, CVE, CAPEC, CWE, flux d'informations suspects, etc.). Une approche prometteuse à l'étude dans ce domaine exploite les techniques d'apprentissage automatique pour automatiser et optimiser les performances des audits de sécurité IoT complexes, et pour améliorer la vitesse et la précision de la détection des incidents de sécurité dans l'IoT. Le traitement des incidents de sécurité au niveau des objets connectés comporte aussi son lot de difficultés, notamment la conception de nouvelles stratégies de confinement des attaques et de mécanismes s'inspirant de paradigmes venant du domaine de la sûreté, tels que *fault-isolation*.

↗ L'équipe-projet **RESIST** d'Inria travaille à la conception de plates-formes de gestion réseau pour l'IoT, et de nouvelles techniques facilitant les fonctionnalités d'audit et de contrôle, afin d'automatiser l'évaluation de la sécurité dans les environnements IoT. On peut citer l'exemple de SCUBA, une plate-forme mise au point afin d'accélérer les audits de sécurité sur des objets connectés hétérogènes.

Sécurisation des logiciels dans l'IoT

Actuellement en plein essor, les logiciels embarqués sur des objets connectés bas de gamme étaient jusque récemment propriétaires et fermés, et parfois même pire : ils reposaient sur le principe de *la sécurité par l'obscurité,* faible par principe. La tendance actuelle est cependant à la mise en œuvre d'un plus grand nombre de logiciels libres pour l'IoT efficace en énergie, – avec pour corollaire l'impossibilité à s'appuyer sur la sécurité par l'obscurité désormais. Avec cette évolution, on peut s'attendre à ce que la sécurité soit donc nécessairement améliorée.

Mais, étant donné le caractère récent de ces implémentations, les vulnérabilités des logiciels de leurs éléments critiques n'ont pas été vérifiées formellement. Pour la communauté des chercheurs dans le domaine de la vérification formelle, le défi à venir est donc d'approfondir l'étude de ces implémentations IoT. L'élaboration de logiciels sécurisés est une tâche ardue, et il n'existe pas d'approche de vérification

universelle. Le principal enjeu consiste à produire des logiciels IoT vérifiés pour les objets connectés bas de gamme, sans compromettre significativement leur performance et leur polyvalence (les logiciels IoT de bas niveau doivent pouvoir s'exécuter sur une grande variété de matériels et d'applications)..

↗ Les équipes-projets **TEA** et **PROSECCO** d'Inria travaillent à l'automatisation des vérifications des briques logicielles intégrables dans des objets connectés de faible puissance. Elles mettent particulièrement l'accent sur la vérification des composants de sécurité critiques, tels que les primitives cryptographiques, et visent à garantir leur exactitude fonctionnelle et/ou la sécurité mémoire. Le développement de nouveaux cadres pour la vérification autour du langage formel *Fstar* facilite la production de modules logiciels embarqués vérifiés et efficaces pour les équipements de faible puissance. La bibliothèque cryptographique HACL est un tel exemple de module logiciel pour l'IoT.

Il est a priori impossible (tant sur le plan technique qu'économique) de vérifier formellement tous les logiciels qui sont vendus et installés. En outre, même si certains logiciels sont formellement vérifiés avant d'être déployés sur le terrain, les logiciels IoT peuvent encore présenter des bugs et des vulnérabilités susceptibles d'être exploités [3]. Ceci s'explique par le fait que le code est contrôlé par rapport à un modèle de sécurité (hypothèses concernant l'attaquant, etc.), et que la vérification préalable n'offre aucune garantie si le modèle ne correspond pas à la réalité. Le modèle de sécurité est vite dépassé quand l'objet connecté est utilisé d'une manière imprévue ou dans un contexte inattendu… et les chances que ce soit le cas ne sont pas minces.

Il s'avère donc nécessaire de compléter la vérification formelle effectuée *a priori* par une mise à jour régulière des logiciels utilisés sur les objets connectés, afin de corriger les bugs et les vulnérabilités qui sont découverts a posteriori, c'est-à-dire après le déploiement du logiciel. Bien qu'il s'agisse d'une fonctionnalité de sécurité, la mise à jour logicielle est également un vecteur d'attaque. Une mise à jour logicielle peut par exemple entremêler un logiciel légitime et un programme malveillant. Autre cas : une mise à jour logicielle opérationnelle et nécessaire peut être bloquée parce qu'aucune partie autorisée n'appose de signature numérique. Un enjeu important dans ce secteur est donc la conception et la sécurisation d'une chaîne d'approvisionnement adaptée pour les logiciels utilisés dans l'IoT, qui doit rester opérationnelle pendant toute la durée de vie des objets connectés

3. Donald Knuth, 1977 : « *Attention aux bogues dans le code ci-dessus ; j'ai seulement prouvé qu'il était correct, je ne l'ai pas testé.* » http://www-cs-faculty.stanford.edu/~knuth/faq.html

(de faible puissance). La recherche sur la cryptographie basse consommation, les logiciels reproductibles, l'attestation à distance et la conception de logiciels systèmes profondément enfouis sont autant de défis connexes. Au-delà de la recherche universitaire, de nouvelles normes sont également nécessaires et souhaitées dans ce domaine, comme le montrent notamment les travaux en cours sur les spécifications SUIT.

↗ L'équipe-projet **TRIBE** d'Inria travaille à la conception d'une chaîne d'approvisionnement sécurisée pour la mise à jour des microprogrammes IoT, adaptée aux appareils IoT de faible puissance peu coûteux, mais sans compromis sur la sécurité.

↗ RIOT-fp est un projet de cybersécurité engagé par Inria, qui cible les objets connectés à base de microcontrôleurs et à ressources limitées. RIOT-fp apporte les briques de base pour concevoir l'IoT open source, en améliorant à la fois la durabilité des logiciels et le compromis fonctionnalité/risque pour les utilisateurs finaux. Ces briques de base combinent des primitives cryptographiques IoT à haute vitesse, haute sécurité et faible impact mémoire, des cadres offrant des garanties pour l'exécution de logiciels sur des objets connectés bas de gamme, et une chaîne d'approvisionnement sécurisée pour les mises à jour logicielles dans l'IoT sur des réseaux à basse consommation énergétique.

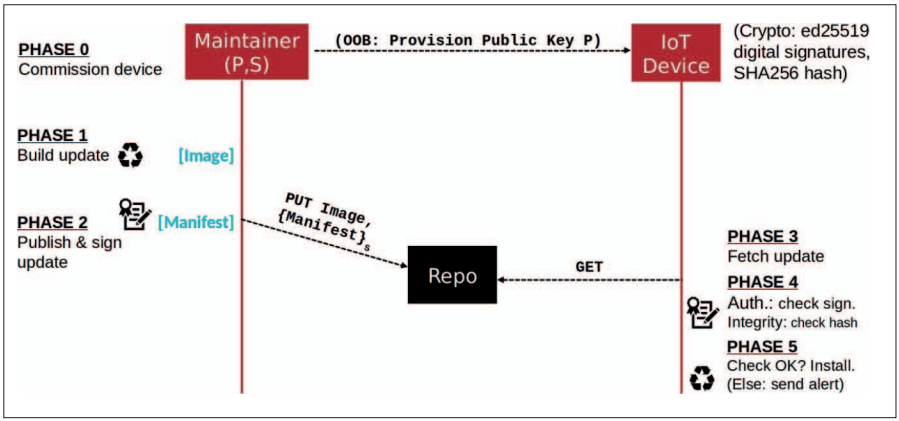


Figure 2 : Interactions lors d'une mise à jour logicielle sécurisée dans l'IoT (spécification SUIT en cours de développement)..

Sécurisation du matériel dans l'IoT

La flexibilité inhérente aux applications IoT fonctionnant dans un environnement multinorme exige une interopérabilité, une facilité d'utilisation et des fonctions de mise à jour des produits. Ces caractéristiques ne sont généralement prévues que dans le développement de logiciels de haut niveau qui contrastent avec le logiciel bas niveau s'exécutant sur les objets connectés. L'accélération matérielle peut augmenter l'efficacité énergétique de plusieurs ordres de grandeur, là où les approches purement logicielles sont souvent incompatibles avec les contraintes de ressources des objets connectés.

D'un côté, une conception hybride intelligente - avec une architecture mixte matériel/logiciel – constitue donc une piste prometteuse qui doit être explorée plus avant. Parmi les exemples de fonctionnalités critiques pour lesquelles une conception hybride est nécessaire, on peut citer (de façon non limitative) la cryptographie IoT.

D'un autre côté, le matériel offre également une surface d'attaque qui doit être contrôlée par des mécanismes spécifiques. Dans les objets connectés de faible puissance en particulier, l'accès physique et la sécurité des appareils doivent être réévalués. Au lieu de pirater votre *smartphone* ou votre ordinateur portable, il peut être plus facile pour un attaquant de pirater un objet connecté (par exemple un capteur/actionneur parmi les dizaines disséminés dans les alentours) et de soumettre ce matériel à des attaques par canaux auxiliaires élaborées. Dans ce contexte, les enjeux consistent à :

- disposer de nouveaux accélérateurs matériels pour les fonctions de sécurité (cryptographies symétrique et asymétrique, hachage, authentification, signature, génération de nombres aléatoires, etc.), en portant une attention particulière à l'ultra-basse consommation énergétique ;
- développer des cryptoprocesseurs spécialisés incluant une protection contre les attaques comme la randomisation ;
- optimiser des compilateurs ciblant les cryptoprocesseurs à ressources limitées ;
- utiliser la traduction binaire dynamique (DBT) à accélération matérielle pour améliorer la protection des logiciels ;
- concevoir de nouvelles techniques permettant de protéger efficacement le matériel contre les attaques par canaux latéraux et l'injection de fautes, en utilisant une approche jointe logicielle-matérielle..

↗ L'équipe-projet **CAIRN** d'Inria travaille sur l'accélération matérielle des primitives cryptographiques à basse consommation énergétique, ainsi que sur des architectures de cryptoprocesseurs peu énergivores comportant des contre-mesures de sécurité matérielles. L'équipe-projet **PACAP** étudie des mécanismes de sécurité insérés par le compilateur pour améliorer la productivité des programmeurs et la robustesse des applications contre les attaques par canaux auxiliaires.

2.10 Architecture matérielle de faible puissance, programmation et compilation

Comme déjà mentionné dans le présent Livre blanc, l'un des enjeux transversaux les plus importants est l'efficacité énergétique et, de manière plus générale, l'efficacité en terme de consommation de ressources non renouvelables. De nombreux objets connectés doivent fonctionner pendant des années avec une petite batterie qui n'est pas destinée à être changée ou rechargée. L'enjeu consiste d'une part à mieux concevoir l'alimentation les objets connectés en énergie, d'autre part à réduire leur consommation d'énergie, tant sur le plan matériel que logiciel. En outre, d'un point de vue plus global, les objets connectés devraient se compter en milliards. Or, même une petite diminution de la consommation d'énergie individuelle de ces milliards de dispositifs peut permettre d'économiser d'importantes quantités d'énergie, de réduire considérablement les coûts et de diminuer l'impact environnemental. Il convient donc d'explorer toute une série de pistes de recherche complémentaires. Quelques-unes sont détaillées ci-après.

De l'ultra-basse consommation au bilan énergétique nul

Pour des raisons liées autant au facteur de forme qu'au coût ou à la maintenance, il est peu pratique d'utiliser des câbles, voire des batteries, pour alimenter les systèmes embarqués. Certains travaux de recherche ont donc pour but de concevoir des dispositifs communicants autonomes, qui ne sont alimentés ni par batterie ni par câble : nous parlons ici d'objets connectés « net zéro ».

On peut par exemple citer la radio-identification (ou RFID) passive, qui utilise des tags (étiquettes électroniques) sans batterie. Elles comprennent uniquement une puce et une antenne, et sont alimentées au moment de la lecture des données par un lecteur. Cette technologie est donc très extensible sur le plan énergétique : un seul lecteur peut être utilisé pour un nombre infini d'étiquettes. Les applications activées par la RFID passive sont cependant limitées, car les étiquettes ne peuvent pas communiquer si le lecteur ne se trouve pas à proximité, et le lecteur consomme généralement plus d'énergie qu'un objet connecté basique (en mode actif).

Certaines recherches se concentrent sur le développement de nouveaux équipements IoT sans batterie qui récoltent l'énergie ambiante, par exemple la lumière, la chaleur, les vibrations/mouvements ou les ondes radio. Un tel procédé fournit toutefois des niveaux de courant très faibles. Il faut donc mettre au point des objets connectés « net zéro » permettant de consommer le moins d'énergie possible.

Des techniques de contrôle de la puissance (ou *"power gating"*) permettent de couper le courant au niveau des blocs du circuit qui ne sont pas utilisés. De même, l'utilisation d'une mémoire non volatile (ou NVM) comme la mémoire RAM non volatile (NVRAM) permet théoriquement à un appareil de subir des coupures de courant sans perte de données et de poursuivre sa tâche sans avoir à la redémarrer du début. Les technologies NVM actuelles impliquent des processus d'écriture lents, énergivores, avec une endurance limitée. Dans ce domaine, les enjeux de la recherche reposent donc sur la conception de nouveaux matériels capables non seulement de récolter plus efficacement l'énergie ambiante mais aussi d'offrir de meilleurs niveaux de fonctionnement avec le très faible courant dérivé.

Il est également important de travailler à la conception de logiciels et de protocoles réseau performants et robustes pouvant être exécutés sur ce type de matériel. Concernant les chaînes d'outils utilisées dans ce but, l'un des défis consiste à concevoir des compilateurs capables, dans le cadre de la programmation de dispositifs embarqués alimentés par intermittence, de mieux assister les programmeurs par le biais d'une analyse des programmes. Cette analyse des programmes est indispensable pour identifier des stratégies de *checkpointing* efficaces, c'est-à-dire déterminer quel état de programme doit être mis en mémoire et quand.

↗ L'équipe-projet **PACAP** d'Inria travaille sur des compilateurs et des analyses de programmes conçus pour faciliter le checkpointing sur des systèmes alimentés par intermittence.

L'architecture des logiciels embarqués pose quant à elle d'autres défis.

Le remplacement de la RAM traditionnelle par la RAM non volatile a, par exemple, des effets secondaires indésirables sur le système embarqué. Les interruptions étant fréquentes (pour mise en veille), elles peuvent survenir au cours de la modification d'une structure de données non volatile. Et lorsque la plate-forme se réinitialise, le programme redémarre alors avec des données incohérentes. C'est ce qu'on appelle le problème de la « machine à remonter le temps cassée ». En fait, il est susceptible de survenir sauf si tous les bits de tous les éléments mémoire d'un système (c'est-à-dire l'unité centrale et la mémoire, mais aussi les

périphériques !) sont rendus nonvolatils. Toutes les couches logicielles sont donc impactées par ces choix architecturaux.

Les enjeux principaux dans ce cadre consistent à

- assurer la cohérence des données, en évitant les pertes de performance excessives, ce qui implique de concevoir de nouvelles techniques d'exécution et de compilation ;
- concevoir un système multitâche performant dans un contexte où les coupures de courant sont fréquentes ;
- concevoir des protocoles réseau exploitant l'énergie ambiante, tout en évitant une perte de performance trop importante lorsque les nœuds du réseau se réinitialisent très fréquemment ;
- assurer un service ininterrompu face à une connectivité et/ou une alimentation intermittente(s) des objets connectés.

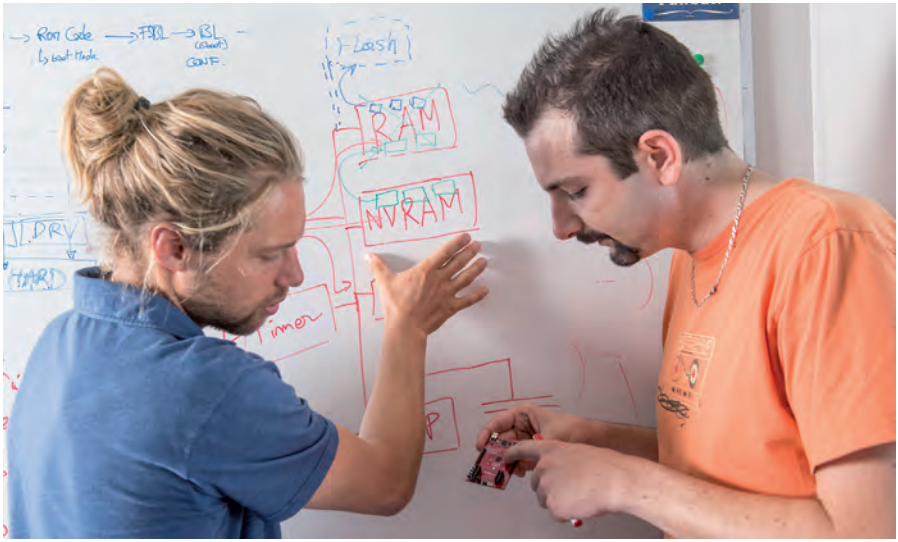


Schéma d'une carte d'expérimentation comprenant un micro-contrôleur avec NVRAM.
© Inria / Photo C. Morel.

↗ L'équipe-projet **SOCRATE** d'Inria travaille à la conception d'architectures logicielles embarquées robustes compatibles avec la récupération d'énergie et l'alimentation intermittente sur des objets connectés « net zéro ».

↗ ZEP est un projet de recherche interdisciplinaire lancé par Inria pour la conception de minuscules objets connectés sans fil et sans batterie, qui récoltent l'énergie dans leur environnement et s'appuient sur une nouvelle architecture à base de RAM non volatile (NVRAM). Pour pouvoir tirer profit des innovations matérielles liées à la récolte d'énergie et à la RAM non volatile, et pour optimiser l'utilisation de l'énergie, ZEP vise à concevoir de nouveaux mécanismes logiciels, actifs au moment de la compilation d'une part, et au moment de l'exécution d'autre part, en y associant des travaux sur l'architecture, la compilation et les systèmes d'exploitation.

Conception d'équipements plus rapides, plus petits et moins coûteux pour l'IoT

L'évolution des technologies matérielles appliquées aux processeurs a atteint une limite. Désormais, la spécialisation du matériel sera sans doute la technique la plus pertinente pour augmenter l'efficacité énergétique (le nombre de calculs par unité de temps et par watt consommé). Les accélérateurs matériels dédiés peuvent permettre de multiplier l'efficacité énergétique par cent (voire plus) comparé aux ordinateurs génériques. Cette augmentation provient essentiellement du transfert des données plus près de la zone de calcul et de la suppression du coût énergétique lié à la programmabilité complète (extraction d'instructions, cache, spéculation, etc.). Les petits appareils embarqués nécessitent un matériel spécifique pour pouvoir fonctionner avec des contraintes strictes en termes de puissance et d'énergie. On s'attend à ce que les accélérateurs matériels deviennent chose encore plus courante dans les dix prochaines années, pour être en mesure de satisfaire les exigences de performance croissantes.

La forte demande visant à pousser (ou à conserver) plus d'intelligence à la périphérie du réseau va renforcer le besoin d'objets connectés plus « éconergétiques ». L'apprentissage automatique et les moteurs d'inférence, par exemple, fonctionnent aujourd'hui via un centre de données distant. Au lieu de cela, la prochaine génération de réseaux neuronaux sera déployée à la périphérie, pour tirer parti des capteurs en temps réel qui collectent les données d'apprentissage, et pour limiter le coût énergétique du transfert des données brutes sur le réseau. Il est donc souhaitable de concevoir des accélérateurs matériels adaptés aux réseaux neuronaux sur les objets connectés de faible puissance.

Cependant, plus le matériel est spécialisé, plus il est difficile à « programmer ». Les concepteurs d'accélérateurs matériels pour les objets connectés, qui

fonctionnent dans la gamme des milliwatts (ou moins), sont confrontés à un certain nombre de défis. Ils doivent explorer un très vaste espace de conception qui englobe à la fois le matériel et les logiciels. Dans ce domaine, les enjeux pour la recherche consistent à :

- identifier et définir les fonctionnalités de la pile logicielle IoT qui se prêtent à l'accélération, les autres étant conservées sur des processeurs génériques programmables à basse consommation ;
- concevoir des interfaces de programmation dédiées pour les accélérateurs, permettant une communication transparente entre logiciel et matériel ;
- assurer la « reconfigurabilité » des accélérateurs en cours d'exécution, avec maintien de la performance ;
- offrir un niveau de programmabilité suffisant dans l'accélérateur (en gros, l'équivalent du GPU pour l'IoT) afin de s'adapter à l'évolution des normes ou des usages..

↗ L'équipe-projet **CAIRN** d'Inria travaille à la conception de plates-formes matérielles de calcul à ultra basse consommation pour l'IoT, spécialisées mais programmables, et de nouveaux niveaux d'abstraction pour des accélérateurs matériels spécifiques. L'équipe-projet **CAIRN** travaille également à l'optimisation des architectures matérielles de processeurs embarqués avec des mécanismes réunissant RAM non volatile et alimentation intermittente.

Vers des objets connectés à l'échelle millimétrique

Les développements récents dans le domaine de la microélectronique ont conduit à l'apparition des premiers prototypes de micropuces d'une dimension inférieure à un grain de riz, qui sont capables de détecter, de calculer et de communiquer sans nécessiter d'autres composants : pas besoin, notamment, de circuit imprimé ni de connexion *"multi-die"*. Cette miniaturisation extrême est notamment rendue possible en supprimant les oscillateurs à cristal externes, et en faisant appel seulement à des circuits oscillateurs de type RC internes, situés à l'intérieur de la puce. Cette micropuce, qui peut être de très petite taille et peu coûteuse, constitue une avancée technologique fascinante. Des défis importants restent toutefois à relever concernant la synchronisation des mécanismes d'horloge sur de tels dispositifs.

L'inconvénient de l'utilisation d'un oscillateur RC, en particulier, est la dérive de l'horloge, qui est de l'ordre de 16 000 ppm pour une micropuce sans cristal,

contre 40 ppm pour un oscillateur à cristal. La dérive est également très sensible à la température. Une micropuce sans cristal nécessite un calibrage manuel minutieux de ses horloges, afin d'assurer une communication réussie avec d'autres dispositifs disponibles sur le marché. L'absence de cristaux dans les micropuces remet en question le fondement même de la recherche sur les communications sans fil de faible puissance. Pratiquement toutes les plates-formes sans fil de faible puissance sont équipées d'un système radio qui communique en toute fiabilité, et est capable de mesurer le temps avec précision. L'apparition récente d'objets connectés de dimension millimétrique exempts de cristaux ouvre donc un nouveau champ de recherche.

↗ L'équipe-projet **EVA** d'Inria développe des algorithmes et des protocoles de calibrage permettant aux objets connectés de taille millimétrique de communiquer avec des dispositifs du commerce et de former des réseaux coordonnés. En partenariat avec l'université de Californie à Berkeley, l'équipe-projet **EVA** a mis au point *SCuM*, la première micropuce au monde à satisfaire aux normes de communication sans fil.

Dompter le polymorphisme du matériel de faible puissance

Le matériel informatique standard dInternet a majoritairement convergé vers une configuration quasi universelle combinant des processeurs 64 bits (*x86 ou ARM*).

En comparaison, **la diversité du matériel pour l'IoT de faible puissance est extrême.** Les architectures de processeurs rencontrées sont très variables (de 8 bits à 16 bits, 32 bits et 64 bits), et proviennent d'une grande variété de constructeurs.

Cette très grande diversité matérielle constitue un défi technique en soi : **il est difficile de choisir le matériel adapté, et le développement de logiciels pour l'IoT requiert trop souvent des compétences « exotiques ».** Les problèmes d'interopérabilité sont exacerbés.

L'innovation dans le matériel de faible puissance ne connaît aucun repos, avec un rythme qui ne faiblit pas. Exemple récent : la famille d'architectures de processeur très polymorphe de type *RISC-V*, qui s'apprête à bouleverser le *statu quo* en rivalisant avec les processeurs ARM Cortex-M, dont la domination allait croissant. Côté radio, de nouvelles catégories de puces auto-alimentées (sans batterie) voient le jour, et sont appelées à bouleverser la notion même de faible puissance. Dans le même temps, la miniaturisation extrême promet l'émergence

d'une nouvelle génération de « systèmes sur une puce » (ou *System-on-Chip*, SoC), qui, par leur taille, ressembleront à de la « poussière intelligente ».

Dans ce domaine, le défi consiste donc, dans un premier temps, à exploiter cette extrême diversité, puis à **évoluer vers un nombre réduit de plates-formes matérielles génériques standard de faible puissance.**

Des plates-formes logicielles embarquées standards pour les objets connectés de faible puissance

Les exigences dues aux objets connectés en matière de fonctionnalités à basse consommation, de cybersécurité, d'interopérabilité et de gestion, augmentent considérablement la complexité des logiciels IoT embarqués, y compris sur les dispositifs de faible puissance à base de microcontrôleurs. Auparavant, les logiciels embarqués sur ces dispositifs étaient généralement monotâches, immuables, propriétaires et très spécifiques au matériel et/ou au fournisseur. Ces caractéristiques évoluent à mesure que la complexité des logiciels IoT augmente. On s'attend désormais à ce qu'une grande partie de ces logiciels imite la dynamique des logiciels types de l'ère internet : plus polyvalents, *open source*, réutilisables sur la palette hétérogène de matériels et de fournisseurs, et capables d'implémenter un ensemble de normes et d'interfaces de programmation (API) courantes. Il est devenu nécessaire de favoriser le développement de logiciels IoT génériques dans tous les secteurs de l'industrie (par exemple, un même algorithme de contrôle et une même implémentation seraient appliqués dans différentes industries). Cette évolution a conduit à l'émergencede pléthore de systèmes d'exploitation embarqués qui visent à fournir une plate-forme logicielle adaptée. De nombreux fournisseurs et géants de la technologie de l'information proposent leurs propres plates-formes, soulignant d'une part que celles-ci sont fondamentales et d'autre part qu'il faut s'attendre à une consolidation du marché. Le défi pour ces plates-formes logicielles embarquées consiste à équilibrer les performances (consommation d'énergie et latence ultra-faibles, empreinte mémoire infime, etc.), avec des garanties de sûreté et de sécurité, tout en facilitant le développement et la portabilité du code profondément embarqué sur des matériels IoT de faible puissance extrêmement divers.

↗ Chez Inria, des équipes-projets dont **TRIBE** et **EVA** travaillent à la conception de plates-formes logicielles embarquées compactes et de faible puissance pour les objets connectés, à l'exemple du système d'exploitation RIOT. Un autre exemple est *OpenWSN*, la pile de protocoles réseau 6TiSCH open source de référence.

↗ RIOT est un système d'exploitation universel, destiné aux petits objets connectés qui ne peuvent pas utiliser Linux en raison de contraintes de ressources matérielles. RIOT offre une plate-forme gratuite et open source, développée par une communauté internationale regroupant des entreprises, des universitaires et des férus d'informatique, qui a été cofondée par Inria. L'objectif de cette plate-forme est d'implémenter et de regrouper les briques de base nécessaires pour créer un Internet des objets efficace en énergie, plus durable, sécurisé, transparent et respectueux de la vie privée.

2.11 Optimisation de l'empreinte ressources globale

D'une part, la notion de performance comprend de multiples facettes : vitesse, précision, garanties de sécurité, équité entre utilisateurs, etc. D'autre part, la performance ne s'évalue pas seulement à une échelle locale, mais aussi dans le contexte plus large d'un système global. Face à des contraintes strictes de frugalité qui concernent les ressources locales des objets connectés, de nouveaux compromis doivent être explorés. Plus globalement, face à une crise écologique mondiale, il est nécessaire de procéder à une évaluation complète du rapport empreinte/avantages de l'IoT.

De nouveaux compromis en termes de performance versus consommation énergétique

Aujourd'hui, la plupart des opérations informatiques sont exécutées dans un environnement largement surdimensionné en termes de qualité du résultat (de précision notamment). Dans de nombreux cas toutefois, des résultats acceptables peuvent être produits sur la base de calculs inexacts ou approximatifs. Les applications traditionnelles (signaux, images, vision, communications sans fil, etc.) comme les applications émergentes (apprentissage automatique, exploration de données, etc.) présentent une résilience inhérente aux erreurs. Moins de performance pour une consommation énergétique réduite : c'est un compromis traditionnel, qui doit être revu et adapté dans l'IoT.

Une approche prometteuse est ainsi d'exploiter le compromis énergie *vs* précision (tout en maintenant la fonctionnalité dans des limites acceptables) pour améliorer l'efficacité énergétique, en complément de l'accélération matérielle. On peut par exemple augmenter de plus de cinquante fois l'efficacité énergétique en remplaçant une opération en virgule flottante double précision 64 bits nécessaire pour des calculs scientifiques de haute précision, par une opération 8 bits de faible précision adaptée à la vision (en tenant compte du stockage, du transport et du calcul des données). De même, d'énormes gains peuvent être obtenus en quantifiant strictement les poids d'un modèle d'apprentissage automatique,

afin de s'adapter à la mémoire minuscule et aux petites capacités du processeur disponibles sur un objet connecté de faible puissance.

Jusqu'à présent, les optimisations se sont principalement concentrées sur les représentations de bas niveau du calcul arithmétique, qui ne s'appliquent pas aux applications IoT à grande échelle. L'enjeu consiste maintenant à concevoir des niveaux d'abstraction plus élevés pour améliorer l'évolutivité et à identifier les transformations de haut niveau qui peuvent jouer sur la précision. L'acceptabilité de l'approximation repose sur des connaissances spécifiques à chaque domaine applicatif, qui doivent régulièrement être mises à jour (elles peuvent évoluer) et être exploitées efficacement. Le degré d'approximation peut être ajusté par le programmeur au moment de la conception ou au moment de l'exécution. Intégrer l'analyse du compilateur et les transformations dans l'ajustement de la précision est un défi (notamment l'identification des zones prometteuses, la décomposition hiérarchique des grands programmes et les transformations algorithmiques).

Sur la base de principes similaires, l'exploration d'autres compromis, comme la vitesse par rapport à la consommation d'énergie, ou l'amélioration de la sécurité par rapport à la consommation d'énergie, constitue un autre défi pour les chercheurs.

↗ L'équipe-projet **CAIRN** d'Inria étudie les compromis entre la précision et la consommation d'énergie, en mettant au point des méthodes pour optimiser les architectures de calcul de faible précision destinées à l'IoT, ainsi que l'inférence et l'apprentissage pour les réseaux neuronaux profonds placés à la périphérie du réseau. L'équipe-projet **TRIBE** étudie les compromis en termes de coûts de communication et de calcul, de précision et de confidentialité, avec des approches basées sur des modèles d'apprentissage automatique hiérarchiques, qui visent à partager et à distribuer l'inférence le long du continuum IoT, depuis le dispositif contraint jusqu'au cloud, en passant par la périphérie.

Les objets connectés consomment de l'énergie non seulement pour détecter, calculer et traiter les données, mais aussi pour communiquer sur le réseau. Une approche prometteuse pour réduire la consommation énergétique consiste ainsi à réduire la fréquence et la taille des transmissions de données. Cependant, le fait d'envoyer moins de données peut diminuer la précision de celles disponibles à distance. Il y donc aussi un compromis à faire entre précision des données et « élagage » des communications.

Dans ce domaine, un champ de recherche se concentre sur la conception de mécanismes de double prédiction, qui utilisent des techniques d'apprentissage

automatique pour déduire la prochaine transmission de données qui est nécessaire, sur la base des transmissions précédentes. Les nouvelles données ne sont transmises que si la prédiction diffère trop fortement des nouvelles données. L'enjeu consiste ici à adapter les moteurs d'inférence et les modèles d'apprentissage automatique aux budgets ressources (mémoire/calcul) extrêmement réduits dont disposent les objets connectés.

Continuité du service en présence d'une connectivité ou d'une alimentation intermittentes

Afin de réduire la consommation globale des objets connectés, des recherches sont menées sur la conception de la pile de protocoles de communication, afin de définir des cycles d'utilisation tels que chaque nœud puisse couper régulièrement sa (ses) interface(s) de communication. On dit alors que le nœud est en sommeil. En effet, la communication est une tâche des plus énergivores pour un nœud IoT (comparé à la détection et au traitement). Mais lorsque l'interface de communication est désactivée, le nœud est déconnecté et ne peut recevoir aucun message. Si un message est envoyé à un nœud en sommeil, le message sera perdu et l'énergie utilisée pour l'envoyer gaspillée.

Dans un système IoT traditionnel, plusieurs mécanismes peuvent être utilisés pour indiquer à un nœud, lorsqu'il se réveille, qu'il devrait rester éveillé plus longtemps. Par exemple, le système pourrait être entièrement synchronisé, et les nœuds sauraient ainsi quand ils doivent se réveiller et écouter, et quand ils peuvent se mettre en sommeil. Cependant, obtenir une synchronisation précise dans des réseaux fortement distribués est déjà un défi en soi ! *A contrario*, dans les réseaux asynchrones l'expéditeur doit, de manière générale, envoyer régulièrement soit un bref signal d'alarme, soit la totalité des données jusqu'à ce que le récepteur se réveille, reçoive le signal d'alarme ou les données et accuse réception auprès de l'expéditeur.

Le matériel des appareils IoT propose généralement des modes d'économie d'énergie (modes veille) qui consomment une énergie négligeable, mais qui nécessitent de désactiver temporairement l'unité centrale et les interfaces réseau. Néanmoins, pour une communication réseau réussie, il faut que l'émetteur et le récepteur soient actifs en même temps. Le compromis consiste alors à permettre à chaque dispositif de dormir autant que possible, tout en l'activant aux moments appropriés pour assurer une fonctionnalité globale.

Les réseaux IoT synchrones résolvent ce problème en programmant à l'avance les moments où les dispositifs se réveillent et écoutent, et ceux où ils sont en sommeil. Dans ce contexte, c'est un défi ardu de concevoir des mécanismes de planification intelligente et de synchronisation précise avec un coût réduit dans des réseaux fortement distribués.

Les réseaux IoT asynchrones (qui représentent jusqu'ici la majeure partie de l'IoT) présentent d'autres défis. Dans ce domaine, la conception de « radio de réveil » fait l'objet de travaux de recherche actifs. Avec cette approche, les appareils sont équipés de deux interfaces radio. L'interface principale, utilisée pour le transfert de données, est désactivée par défaut. Une interface secondaire, de très faible puissance, est utilisée pour recevoir les signaux de réveil. Des récepteurs de réveil passifs sont à l'étude, l'enjeu étant d'augmenter leur sensibilité sans augmenter la puissance de transmission.

Une approche complémentaire consiste à mettre les données dans une mémoire cache, pour le compte des nœuds en sommeil, quelque part dans le continuum *Cloud-Edge*-Objet. L'enjeu consiste alors à déterminer et à évaluer de nouvelles stratégies de mise en cache et de remplacement du cache qui sont susceptibles :

- d'optimiser le lieu où les données IoT doivent être stockées : le coût est réduit lorsque cet endroit se situe plus près du demandeur, mais il peut alors se trouver loin de la source des données ;
- d'optimiser le nombre de copies à stocker : les emplacements multiples augmentent le coût de sauvegarde des données, mais peuvent réduire le coût de leur récupération puisqu'ils sont susceptibles d'être plus proches du demandeur ;
- d'optimiser la fréquence des mises à jour : faire davantage de mises à jour améliore la précision, mais augmente le coût.

Évaluer et réduire l'empreinte globale de l'IoT

La crise écologique actuelle pousse les chercheurs de tous les domaines à évaluer l'impact environnemental des différentes technologies, qu'elles soient déjà mises en œuvre aujourd'hui ou à venir. L'IoT permet, entre autres, de mieux trier et recycler les déchets, d'assurer un meilleur éclairage public dans le respect de l'environnement ou de mieux gérer le trafic routier. L'IoT peut donc contribuer

à réduire notre impact sur l'environnement de bien des manières, comme le montrent plusieurs études [4] [5] [6].

Pourtant, les travaux de recherche se concentrent trop souvent sur les optimisations potentielles auxquelles les objets connectés pourraient contribuer, et passent à côté d'une analyse globale, qui permet d'étudier les conditions dans lesquelles des gains nets sont effectivement réalisés (et si ces conditions sont susceptibles d'être remplies ou non).

L'impact environnemental direct doit être évalué en tenant compte du cycle de vie complet des appareils, depuis leur production (par exemple l'extraction de ressources minérales) jusqu'à leur fin de vie (par exemple le recyclage éventuel), en passant par leur coût opérationnel (par exemple la maintenance et la consommation d'énergie). L'amélioration de la recyclabilité des petits composants électroniques utilisés dans les objets connectés constitue en la matière un immense défi pour les chercheurs.

Tous les objets connectés sont constitués de composants électroniques. L'un des enjeux fondamentaux consiste donc à concevoir et à fabriquer des équipements en utilisant moins de ressources. Certains travaux de recherche sont ainsi axés notamment sur la miniaturisation des cartes de circuits imprimés, par exemple pour utiliser moins de métal et de plastique, ou pour permettre l'utilisation de nouveaux matériaux prometteurs tels que le graphène. Les chercheurs ont également certains défis à relever dans la conception d'antennes (qui peuvent être imprimées avec de l'encre biodégradable [7]) et de batteries plus respectueuses des ressources.

Il faut en outre mesurer l'impact indirect, qui englobe les effets induits, les effets de rebond, etc., car certains nouveaux usages sont susceptibles de contrecarrer les optimisations permises par l'IoT. Même si des milliards de dispositifs microscopiques sont utilisés sans fil et sans batterie (alimentés par la récupération d'énergie) pour fournir des services avancés dans la prochaine génération d'Internet, ils peuvent malgré tout contribuer, à l'échelle mondiale, à l'augmentation des émissions de gaz à effet de serre.

4. 5 ways the IoT is helping the Environment
5. Where IoT Meets The Environment: Building a Greener Future
6. IoT for Environmental Sustainability
7. D. Iba et al. *« Development of smart gear system by conductive-ink print »*, Proc. SPIE, 2019.

La recherche se concentre généralement sur l'amélioration de l'impact direct potentiel, mais peine à répondre aux questions plus globales : jusqu'où vont, en pratique, les améliorations ? Jusqu'à quand ? À quel coût ? Et surtout, avec quels bénéfices nets à l'échelle mondiale ? Dans ce domaine, la complexité augmente car des compétences hautement interdisciplinaires sont requises, impliquant des connaissances technologiques, mais aussi sociales, économiques et politiques.

L'un des enjeux en la matière est donc de développer des cadres conceptuels adéquats permettant de mieux appréhender et évaluer l'ensemble des incidences de l'IoT sur l'environnement. Les cadres existants ne prennent généralement pas en compte l'impact indirect, bien que celui-ci puisse éclipser totalement l'impact direct. Même si l'impact direct peut en théorie être mesuré par les cadres existants, ceux-ci sont mis à mal parce que (i) les données sont difficiles à collecter, et (ii) les technologies et les utilisations évoluent à un rythme très rapide.

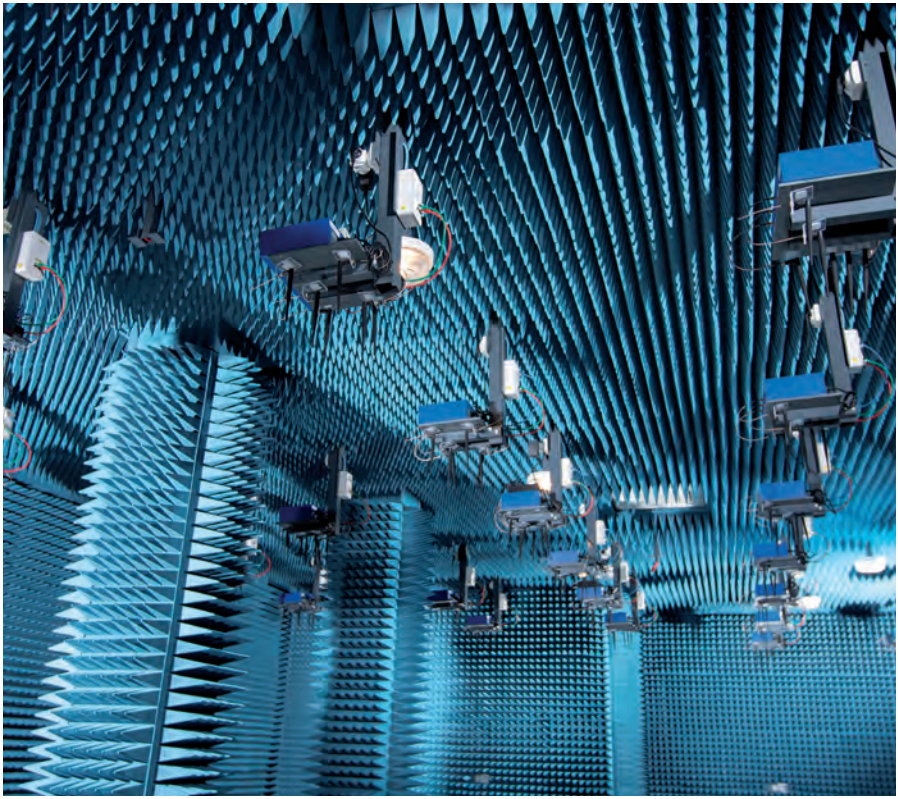
Conclusion

L'Internet des Objets a acquis une importance fondamentale dans le paysage des technologies qui sont appelées à façonner l'avenir. Dans le monde qui s'annonce, les entités (pays, organisations, entreprises, individus) qui souhaitent préserver leur souveraineté doivent devenir plus conscientes, doivent s'engager dans d'importants développements technologiques et des travaux de recherche conséquents dans plusieurs domaines qui sous-tendent l'IoT.

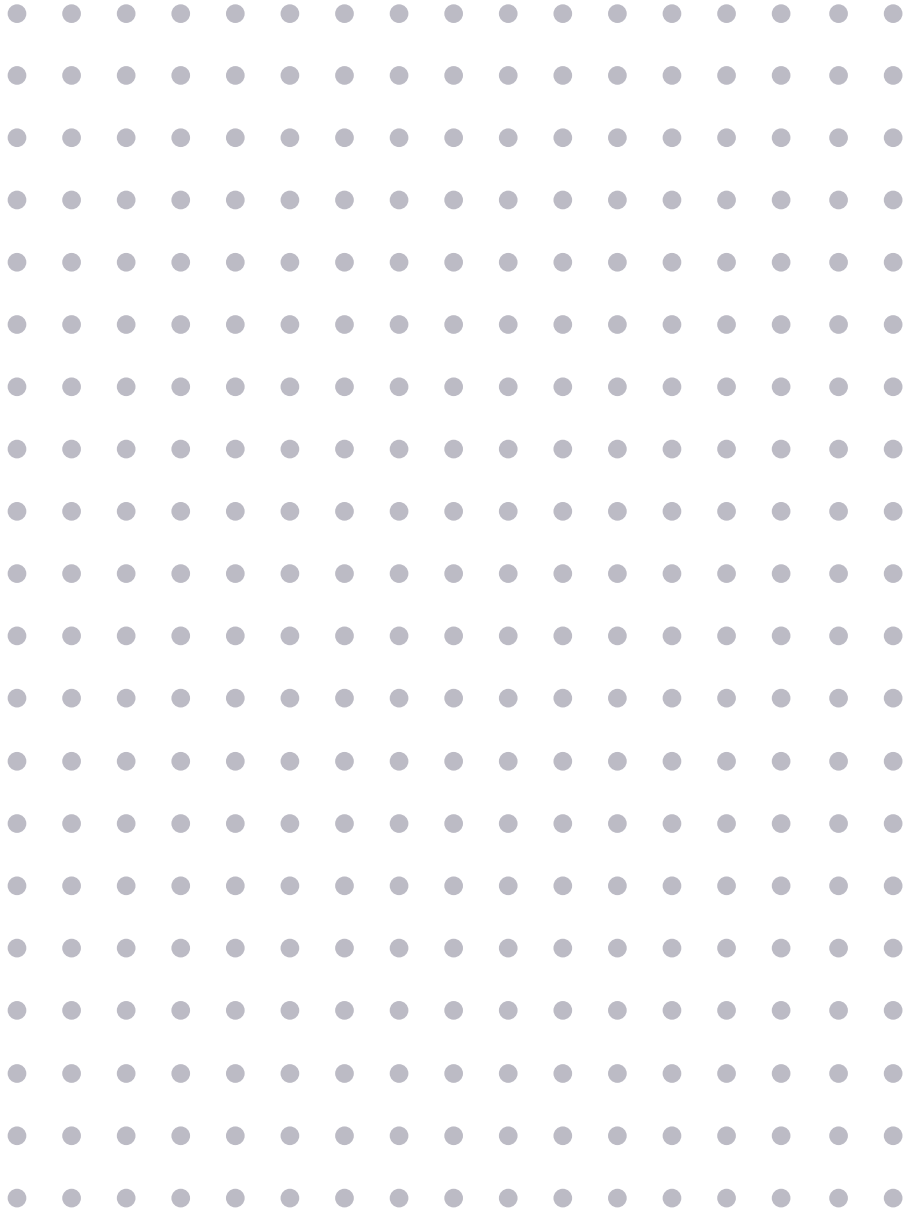
Ces domaines sont variés, allant des réseaux de communication nouvelle génération à l'informatique *"pervasive"* (envahissante, pénétrante), des logiciels pour systèmes embarqués au matériel de faible consommation énergétique, de l'interaction Homme-machine au contrôle et à la résilience des systèmes cyber-physiques, de la cybersécurité et la sûreté au traitement des données respectueux de la vie privée.

En outre, comme la technologie IoT s'imbrique partout dans la société et dans nos vies individuelles, l'élaboration des normes technologiques pour l'IoT est plus que jamais cruciale. Dans ce contexte, pour être en mesure de préserver la neutralité géopolitique des technologies IoT, il est fondamental de participer activement aux travaux des organismes de normalisation concernés.

Enfin, avec l'aggravation de la crise environnementale, on peut espérer que notre impact sur la nature pourra être réduit grâce aux améliorations et aux nouvelles fonctionnalités permises par l'IoT, qui devrait être déployé et utilisé massivement. Des efforts supplémentaires importants sont néanmoins nécessaires pour déterminer si cette réduction pourra, à l'échelle planétaire, dépasser effectivement l'impact sur l'environnement de la production, du déploiement et de la maintenance de ces nouveaux mécanismes connectés, tout au long de leur cycle de vie.



Salle anéchoïde de la plateforme expérimentale FIT (Future Internet of Things). © Inria / Photo C. Morel.



Domaine de Voluceau, Rocquencourt BP 105
78153 Le Chesnay Cedex, France
Tel.: +33 (0)1 39 63 55 11
www.inria.fr