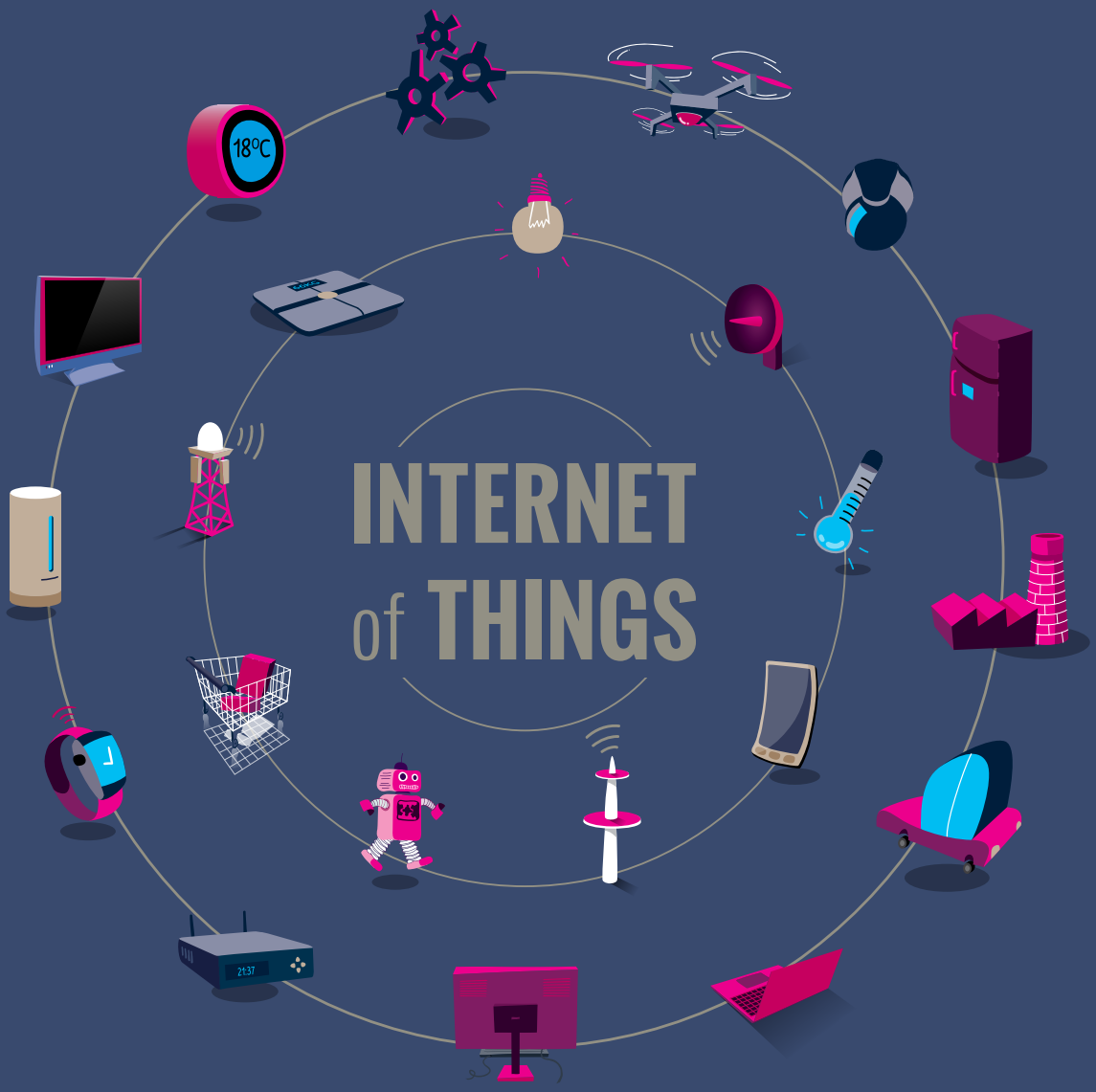
TECH/TRENDS

#11

1ÈRE ÉDITION

PERCEVOIR LE FUTUR



COMMUNIQUER TRAITER SÉCURISER

IOT-EE

TechTrends - Publication de **IoT-ee**

Les TechTrends sont l'expression de notre savoir faire ; forgé sur le terrain, auprès de nos clients dans le cadre des projets que nous menons avec eux.

Fruit d'un travail collaboratif de nos consultants, vous y trouverez, nous l'espérons, les nouvelles tendances technologiques et méthodologiques ainsi que l'état de l'art de notre profession.

Nous tentons, dans le cadre de ces publications, de vous dispenser des conseils directement opérationnels afin de vous guider dans les décisions stratégiques que vous avez à prendre.

Distribués à plusieurs milliers d'exemplaires tous les ans, la collection des TechTrends s'étoffe régulièrement de nouveaux ouvrages.

Luc Legardeur, Président

COMMUNIQUER

TRAITER

SÉCURISER

INTRODUCTION

L'informatique moderne se nourrit constamment de buzzwords. En 2017, Internet des Objets (Internet of Things - IoT) est sur toutes les lèvres.

Le terme IoT est utilisé pour la première fois au laboratoire Auto-ID center au MIT en 1999. Depuis 2013, le dictionnaire OXFORD définit l'IoT comme : ***The interconnection via the Internet of computing devices embedded in everyday objects, enabling them to send and receive data.***

Depuis cette date, le terme IoT fait le buzz et force est de constater que tout est sujet à créer des capteurs pour projeter le monde physique dans le monde numérique. Les objets communicants fleurissent partout : des thermomètres connectés dans les chambres froides jusqu'à la collecte de signes vitaux pour le monitoring médical. La quantité d'objets connectés se multiplie à grande vitesse. Selon Cisco, **ce sont 50 milliards d'objets qui feront partie de notre quotidien d'ici à 2020[1].**

Si l'on omet les progrès de l'électronique en terme de miniaturisation et de baisse des coûts, cette explosion est la conséquence directe de deux avancées majeures dans le monde des télécoms :

- Dans les pays développés, chaque utilisateur possède **au moins un accès vers Internet à travers sa box ou son terminal mobile.** Cela rend possible l'interconnexion des objets de la personne et de la maison sur un réseau local via Bluetooth ou WiFi.
- **Le déploiement des protocoles LPWAN** (Low Power Wide Area Network) est devenu une réalité et les territoires couverts sont de plus en plus étendus, permettant de déployer in-vivo des objets pour une durée de plusieurs années.

1- https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

Il est donc devenu facile d'envoyer des données via divers réseaux vers un Backend. Encore faut-il que ce dernier soit capable de les réceptionner et de les traiter correctement.

Depuis l'avènement du Big Data, **des millions d'objets envoyant des centaines de messages par minute ne représentent plus un défi technologique insurmontable.** De la même façon, de nombreuses problématiques connexes, comme le Machine Learning, la Data Visualization, ou le stockage à grande échelle, ont été largement défrichées.

Ainsi, on peut constater que les architectures IoT n'ont en réalité pas grand chose d'innovant et utilisent des composants rodés au cours de ces dernières années.

Néanmoins, **les défis sont encore immenses tant au niveau de la sécurité des données que de leurs usages,** deux problématiques qui apparaissent comme les parents pauvres des dernières innovations numériques autour de l'IoT.

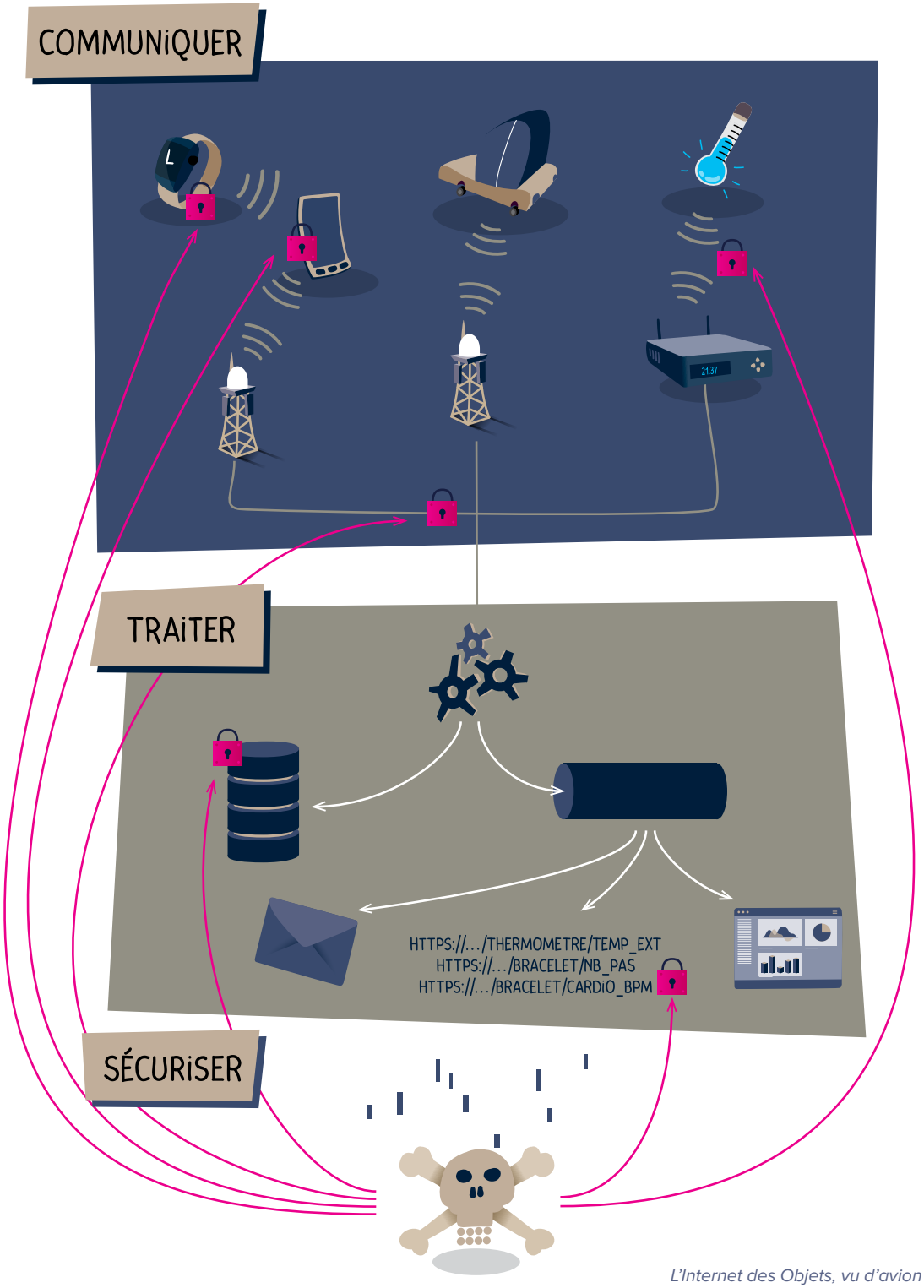
Enfin, et avant tout, l'Internet des Objets concerne ... des objets. Nous ne sommes que des amateurs éclairés dans les arcanes obscures de l'électronique et de la miniaturisation. Néanmoins, aborder le sujet par son axe immatériel ne nous empêche pas de prototyper circuits et objets imprimés dans l'atelier que nous avons monté.

Sachez **qu'il est possible, dès aujourd'hui, d'expérimenter une idée pour un prix raisonnable.** Les composants électroniques sont facilement accessibles. Les scripts permettant d'interroger des capteurs ou d'opérer des actionneurs sont disponibles en open-source dans une multitude de langages. L'émergence des FabLab et de la mouvance DIY (Do It Yourself - Réalisez-le vous-même) met à la portée de tous des machines et des savoir-faire industriels.

Ce TechTrends démystifie l'Internet des Objets et vous ouvre les portes d'un nouveau pan de la transformation numérique des entreprises. Il s'articule autour des 3 actions principales de la chaîne immatérielle de l'IoT :

- **Communiquer :** Comment choisir et connecter vos objets aux différents protocoles IoT disponibles ?
- **Traiter :** Comment collecter, traiter et visualiser vos données IoT ?
- **Sécuriser :** Comment sécuriser l'accès à vos objets et données IoT ?

> *Depuis l'avènement du Big Data, des millions d'objets envoyant des centaines de messages par minute ne présentent plus un défi technologique insurmontable.*



L'Internet des Objets, vu d'avion

01

COMMUNIQUER

Sur le papier tout est prêt.

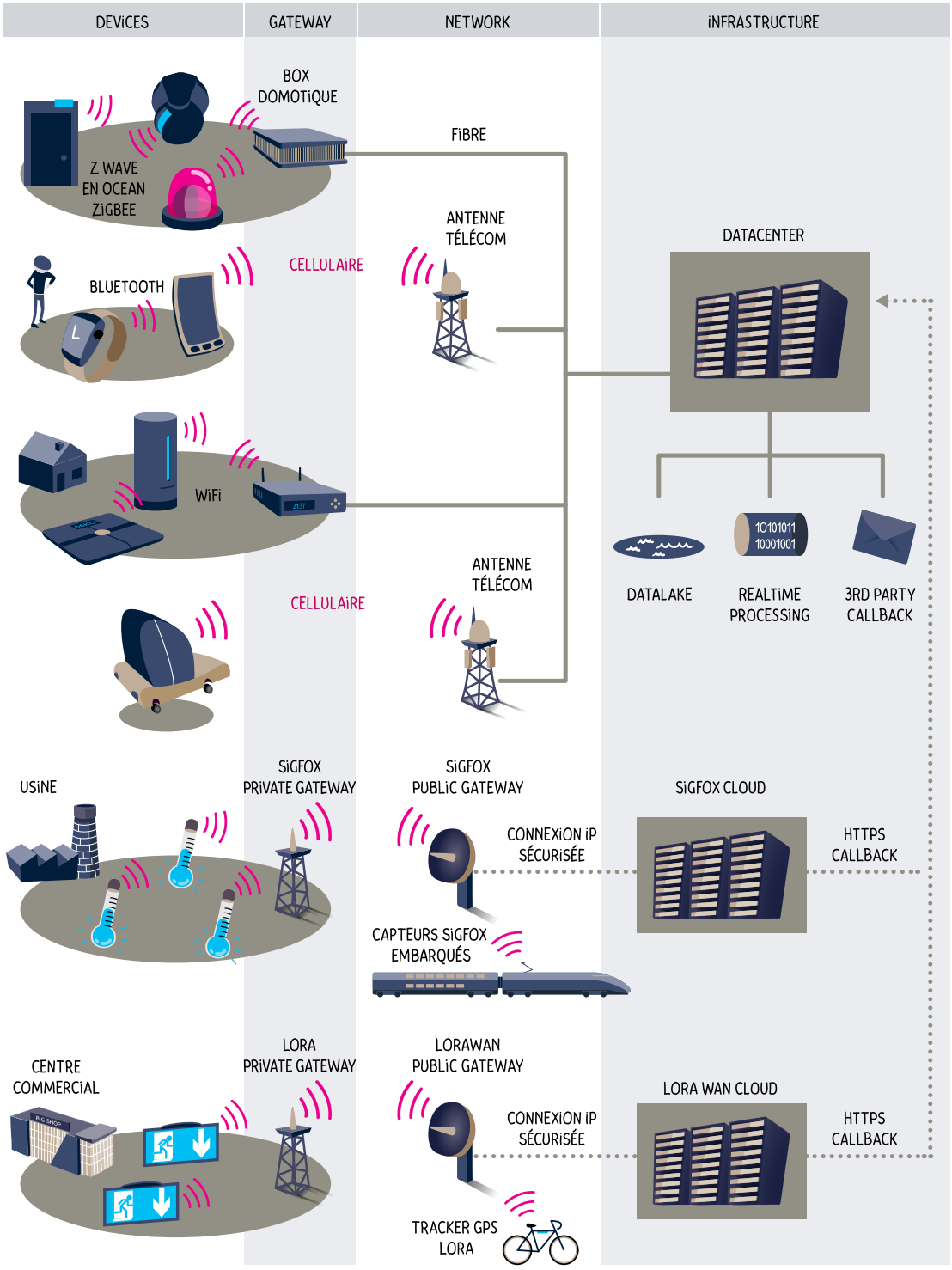
Votre idée est innovante, le Business Plan est monté. Le prototype est fabriqué, il collecte déjà des données.

La première question qui se pose est : **Comment communiquer avec le monde extérieur ?** Faut-il utiliser une boucle locale puis Internet ? Privilégier les réseaux 3G / 4G ? Ou bien simplement se connecter à un réseau WiFi existant ? L'un des problèmes majeurs de l'IoT est qu'il n'existe pas de standards dans les protocoles de communication mais plutôt une multitude de protocoles qui co-existent, comportant chacun ses caractéristiques et ses contraintes propres. Parmi cette multitude de protocoles, citons en vrac NFC, RFID, Bluetooth LE (et très bientôt Bluetooth 5), Sigfox, LoRa, les réseaux cellulaires, etc.

Comment alors s'y retrouver dans une telle jungle ? Quel choix faire pour mon objet et dans quelle mesure ce choix conditionne-t-il l'architecture à venir pour supporter ma flotte d'objets ?

De nombreux réseaux pour différents usages

Les réseaux existants, dédiés à l'Internet des Objets ou non, présentent des caractéristiques techniques très hétérogènes. La multitude de critères existant offre des possibilités de regroupement infinies.



Les différents réseaux utilisés par les objets connectés.

Pour notre analyse, nous avons décidé de nous baser sur la portée et la consommation énergétique des réseaux et dégager ainsi trois grandes familles :

Les réseaux LAN (Local Area Network). Ce sont des réseaux courte portée (de 1 à 100 mètres environ) et peu consommateurs d'énergie. Ces réseaux sont très utilisés comme boucles locales par les objets connectés grand public (à l'image des box domotiques pour la maison ou des bracelets de type fitness). On peut citer par exemple NFC, BLE et Zigbee.

Les réseaux cellulaires. Ce sont des réseaux longue portée (de quelques kilomètres en ville à 30 km en zone rurale) et assez gourmands en énergie. À l'image des réseaux GSM, 2G, 3G ou 4G, ils permettent le transport de grands volumes de données (vidéos, images, etc.) et ont une bonne couverture au niveau national et international (grâce au roaming).

Les réseaux LPWAN (Low Power Wide Area Network). Les réseaux LPWAN se caractérisent par une portée de l'ordre de plusieurs kilomètres et une faible consommation énergétique. À l'image des réseaux Sigfox et LoRa, ils ont été spécifiquement conçus pour transporter les données IoT à faible coût. Pour autant, la multiplication des métriques, le nombre et la durée de vie des capteurs peuvent engendrer des coûts faramineux.

Gérer une flotte de capteurs dont l'activité s'étendra potentiellement sur plusieurs années, c'est aussi prendre en compte un facteur d'échelle. Ces contraintes économiques devront être intégrées au choix des réseaux permettant les échanges.

Les LAN et le concept de Gateway

Dans l'IoT, l'un des principaux enjeux est d'amener la donnée des objets vers Internet. Or, souvent, par design et soucis d'économie d'énergie, les objets ne disposent pas d'une connectivité réseau poussée. Il est alors intéressant d'utiliser des boucles radio locales pour atteindre un « concentrateur » (gateway) qui propagera l'information vers Internet.

La connexion directe vers Internet paraît de plus en plus utopique : Internet n'est pas dimensionné pour gérer l'adressage de milliards d'objets connectés, même avec IPv6.

On peut distinguer deux types de gateway :

- **Les gateways de type box (qu'elles soient domotiques ou internet) :** principalement utilisées pour des applications domotiques, elles reposent sur un grand nombre de protocoles radio courte portée, souvent incompatibles entre eux. On parle par exemple de Zigbee (ampoules connectées Philips Hue), de ZWave (objets de la gamme Fibaro) ou du petit dernier, EnOcean.

L'absence de standard et la concurrence féroce que se livrent les différents acteurs sont autant de freins à l'adoption de la domotique dans les foyers français.

- **Les gateways de type terminal mobile :** aujourd'hui, la majorité des citoyens dispose d'une gateway au fond de leur poche, leur terminal mobile. En effet, ce dernier est pourvu d'une double connectivité très intéressante : une connectivité de courte portée, peu consommatrice d'énergie, via le Bluetooth Low Energy (BLE) et une autre de plus longue portée, via les réseaux 3G / 4G, ou le WiFi, en conditions favorables. Ces caractéristiques en font un relais idéal entre l'environnement de proximité de l'utilisateur et le Web. C'est le mode de fonctionnement de la majorité des bracelets et montres connectés. Il faut néanmoins noter que même si cette stratégie de connexion économise la batterie de l'objet, c'est souvent au détriment de celle du téléphone.

Nous ne détaillerons pas plus le fonctionnement des gateways. En effet, une fois l'information collectée par la boucle locale, les protocoles d'envoi sur Internet sont relativement standard, qu'ils passent par la connexion internet de la maison ou bien par le réseau 3G / 4G du téléphone.

Les réseaux cellulaires

Du côté des réseaux cellulaires, depuis l'apparition de la 4G, il n'y a pas eu de réelle révolution. Les protocoles sont connus et utilisés depuis des années et **les inconvénients sont clairement identifiés, les deux principaux étant la consommation d'énergie et le prix.** En effet, mettre une puce 4G dans tous les objets connectés n'est pas à la portée de toutes les bourses : le coût d'un modem cellulaire 3G / 4G est, en moyenne, d'un ordre de grandeur plus élevé que celui d'un modem LPWAN ; aussi, utiliser les réseaux cellulaires classiques pour transmettre quelques kilo octets par objet et par jour peut engendrer une dépense démesurée quand on passe à l'échelle du déploiement massif.

Les réseaux LPWAN

Pour pallier les défauts de la 4G, les opérateurs télécom ont innové. Ce qui est valable pour une voiture connectée, qui émet en quasi temps réel un nombre incroyable de mesures via un

modem 3G, n'est en aucun cas applicable à une flotte de 10 000 capteurs collectant une tension électrique toutes les heures.

Ceci explique **l'émergence des réseaux sans fil M2M (Machine to Machine) à longue portée, basse énergie et très bas débit.** Les restrictions qu'ils imposent sur la taille des messages se font au bénéfice de tarifs très attractifs.

La France est une nation pionnière dans ce domaine. Nous pouvons nous enorgueillir d'être à l'origine de réseaux comme Sigfox, LoRa, Qowisio ou encore PicoWann.

Les réseaux LPWAN s'appuient souvent sur des bandes de fréquences « ISM » (Industrie / Science / Médical) non licenciées. Ils coexistent sur ces fréquences avec d'autres technologies radio, mais sans risque de collision. Les antennes des émetteurs utilisent la bande des 800-900 MHz et ont une portée plus grande que celles des réseaux 2G / 3G / 4G : elle atteint une quarantaine de kilomètres en zone rurale et quelques kilomètres en ville. De ce fait, le nombre d'antennes nécessaires pour couvrir un pays est de loin inférieur à celui nécessaire pour équiper le même territoire en 4G. À titre de comparaison, Orange a installé 10 000 antennes 4G en France là où Sigfox n'a besoin que de 1 500 antennes pour couvrir une superficie équivalente.

Nous allons nous focaliser sur les deux réseaux phares du moment : **Sigfox et LoRa.**

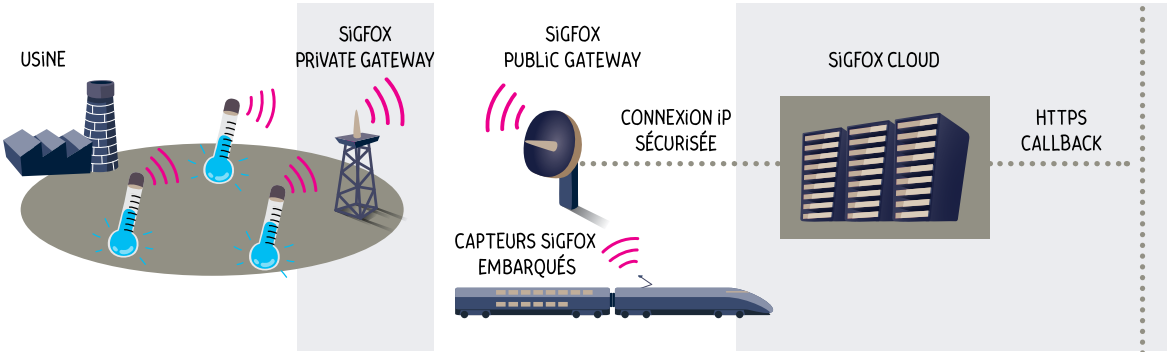
	Opérateurs français	Spécifications	Déploiement France	Déploiement à l'international	Données	Usage
Sigfox	Sigfox	Fermées	100%	Europe du sud, Benelux, Irlande, Autriche, Partiel : US, Brésil, Mexique, Finlande, Allemagne, Japon, Australie	Bidirectionnelle limitée 140 messages/jour 12 octets/message	Envoyer occasionnellement de petites quantités de données.
LoRa	Orange Bouygues Telecom	Ouvertes (Alliance Lora)	En cours	Via des partenaires Comptabilité non garantie Roaming	Bidirectionnelle Pas d'information sur le nombre de messages environ 250 octets / message	Communiquer de manière bidirectionnelle sur des volumes conséquents, en France.

Le réseau Sigfox

Sigfox est un réseau LPWAN qui permet aux objets de communiquer sans fil, entre eux ou avec un concentrateur. Le réseau est fourni par la start-up toulousaine portant le même nom. Elle est l'unique opérateur de son réseau.

Caractéristiques techniques

- **Sigfox repose sur la technologie radio UNB (Ultra Narrow Band)**. Le choix de l'utilisation de bandes étroites permet de faire en sorte que les ondes se propagent sur de plus amples distances.
- Sigfox propose une politique de transfert de données extrêmement frugale : par abonnement, le client est limité à 140 messages par jour, chaque message contenant au maximum 12 octets de données utiles.
- **Sigfox permet une communication bidirectionnelle limitée.** Une telle communication permet par exemple de pousser un paramètre de configuration vers les objets. Cette fonctionnalité est implémentée comme un polling, dans lequel l'objet reste maître et peut demander au Système d'Information s'il y a des données à télécharger. Ceci évite de rester connecté en permanence mais impose que ce soit l'objet qui soit à l'initiative de toutes les transmissions, aussi bien montantes que descendantes.
- **La communication sur Sigfox est sécurisée :** protection anti-rejeu, message de brouillage, séquençage, etc. Seuls les fournisseurs de périphériques comprennent les données échangées entre l'objet et les serveurs d'infrastructure de Sigfox. Ces serveurs agissent ensuite comme un passe-plat, poussant les données vers le système informatique du client sur un canal HTTPS.



Architecture de communication Sigfox

Couverture, modèle économique et application

- Le réseau Sigfox s'étend aujourd'hui sur la majorité du territoire en France, en Espagne, au Royaume-Uni, ainsi que dans certaines villes comme New York ou Moscou. Pour vérifier la couverture dans votre région, reportez-vous à leur catalogue SNO[2].
- Le coût d'un abonnement Sigfox dépend de deux paramètres : le volume de messages échangés par les appareils et le nombre d'appareils. Les prix à l'année varient de 1 à 14€ par appareil.
- Sigfox ambitionne d'attirer les principaux fabricants de l'écosystème. Le protocole Sigfox est supporté par plusieurs composants radio existants comme ceux fournis par Atmel, Silicon Labs, STMicroelectronics, Texas Instruments et Samsung Artik.

Sigfox est un protocole à privilégier pour les applications qui ont besoin d'envoyer occasionnellement de petites quantités de données. On peut envisager son utilisation par exemple pour connecter les systèmes d'alarme ou pour relever à distance des compteurs d'eau ou d'électricité.

Le réseau LoRa

Le CES 2015 marque la création de l'Alliance LoRa[3]. Il s'agit d'un consortium visant à développer **une offre Open Source,** concurrente à celle de Sigfox et basée sur la technologie du même nom brevetée par la société française Cycleo en 2012. Aujourd'hui, l'Alliance comprend 127 membres dont des acteurs majeurs des télécoms français comme Bouygues Telecom, Actility et Sagecomm.

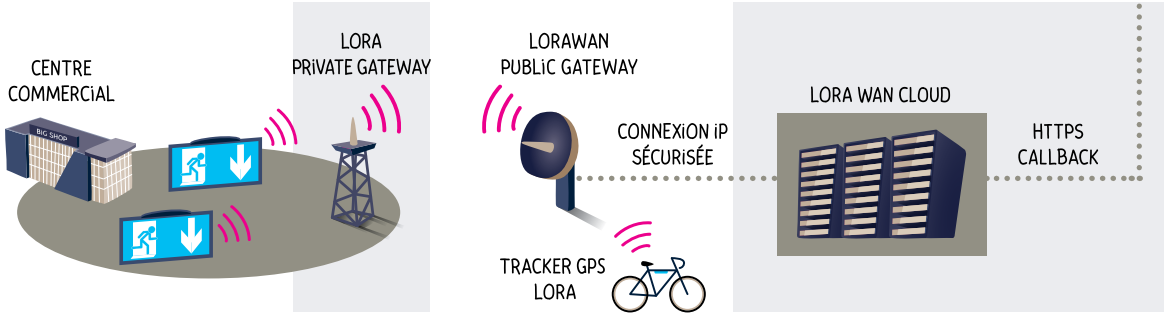
Caractéristiques techniques

- **LoRa est la technologie de modulation à étalement du spectre du protocole LoRaWAN** (Long Range Wide-area ou réseau étendu de longue portée). Si les réseaux de type 3G / 4G se basent sur un protocole IP, le réseau LoRa lui se base sur le protocole LoRaWAN, peu énergivore.

2- http://www.sigfox.com/fr/couverture
3- lora-alliance.org

- **LoRa est une technologie ouverte,** ce qui signifie que n'importe quelle entreprise peut créer son propre réseau LoRa puis l'exploiter. Il faut pour cela se munir d'une antenne reliée à Internet (par Wi-Fi, câble Ethernet, 3G, etc.) avec une station de base émettant en France sur la bande 868 MHz.
- **LoRa est très prisée par l'opérateur Bouygues Telecom.** En effet, ce dernier prévoit de déployer 5 000 à 6 000 stations de base[4] afin de réaliser une couverture de l'ensemble du territoire. L'ambition est même encore plus importante, car la couverture s'étendrait alors au deep indoor, c'est-à-dire jusqu'à un mètre en sous-sol (afin de faire du relevé à distance des compteurs d'énergie par exemple).
- D'après Bouygues Telecom, **les messages envoyés par les objets sous LoRa peuvent aller jusqu'à 243 octets** (cette valeur maximale est non-normée par LoRa et reste donc à la discrétion des opérateurs) mais la moyenne tourne plus autour de 50 à 60 octets. Cependant, plus le message est lourd, plus le temps de transmission est long.
- La portée du réseau LoRa est d'environ 20 km en zone rurale et de 2 km en zone urbaine.

L'architecture du réseau LoRa est très semblable à celle de Sigfox. Le réseau LoRaWAN dispose de cartes électroniques avec des modules LoRa et un / des capteurs implémentés. Ces modules peuvent communiquer entre eux en P2P ou avec une antenne. Les antennes (gateway) reçoivent les données du module radio puis les transfèrent au serveur (backend). Le backend gère le flux de données arrivant des antennes et ces dernières sont directement retransmises en HTTPS, via un callback, au serveur client qui peut les intégrer sur ses applications logicielles.



Architecture de communication Lora et LoraWan

4- corporate.bouyguestelecom.fr/nos-activites/linternet-des-objets

Couverture, modèle économique et application

- Le réseau LoRa est en cours de déploiement et de commercialisation en France. **Bouygues Telecom et Orange espèrent couvrir l'intégralité du territoire métropolitain d'ici mi-2017.** LoRa est présent à l'international via d'autres opérateurs étrangers (USA, Pays Bas, Belgique, etc). Bouygues Telecom et Orange devraient signer des accords avec ces opérateurs pour assurer la continuité du réseau à l'international.
- Pour l'instant, LoRa est uniquement proposée aux clients B2B. Il n'y a pas encore d'informations concernant le prix de l'abonnement. Néamoins Bouygues Telecom et Orange prévoient d'offrir des plateformes Cloud complètes (stockage, traitement) en plus du simple callback, via Objenious (Bouygues Telecom) et DataVenue (Orange).

Comme avec Sigfox, plusieurs applications à faible débit d'informations sont envisageables. On citera par exemple la détection de l'état de fermeture des portes des baraques sur les chantiers, la détection des places libres dans les parking, le relevé des compteurs (eau, électricité, gaz, etc.), la maintenance prédictive, le suivi des approvisionnements et des stocks ou encore le suivi médical des personnes.

> *Le CES 2015 marque la création de l'Alliance LoRa. Il s'agit d'un consortium visant à développer une offre Open Source, concurrente à celle de Sigfox et basée sur la technologie du même nom brevetée par la société française Cycleo en 2012.*

Quel réseau LPWAN choisir ?

Il est difficile de comparer les deux réseaux car certains éléments de comparaison manquent à l'appel, sans oublier le fait que les niveaux de maturité sont eux aussi très différents, LoRa étant toujours en phase de tests.

Les deux réseaux servent certes des marchés similaires mais nous notons dès à présent quelques petites différences qui pourront aider à faire un choix selon votre besoin :

- **Pour envoyer de très faibles volumes de données quelques fois par jour, Sigfox semble être le réseau idéal.** Dès qu'il s'agit de communiquer un peu plus souvent ou d'avoir une bonne bi-directionnalité (beaucoup d'allers / retours entre l'objet connecté et le backend), LoRa semble être le plus adapté. Cependant, les échanges trop nombreux peuvent drastiquement diminuer la durée de vie de la batterie des capteurs.
- Sigfox propose une macro-géolocalisation de l'ordre du kilomètre pour suivre par exemple la logistique des conteneurs. Pour sa part **LoRa, propose une géolocalisation théorique plus fine de l'ordre du mètre.** Nous disons théorique car cette finesse nécessiterait de densifier le réseau et d'augmenter le prix du service proposé.
- Aujourd'hui, **seul Sigfox a une couverture totale de la métropole française.** LoRa est encore en cours de déploiement et de test. Pour une connexion immédiate, le seul choix viable est donc Sigfox. De plus, Sigfox étant déjà déployé dans plusieurs pays étrangers, les utilisateurs peuvent profiter de ces antennes hors de France avec un abonnement standard.
- Même si LoRa est encore en phase de test, **il semble avoir de meilleures performances en termes de couverture indoor et de mobilité,** tout en gardant à l'esprit que ces performances dépendent grandement des designs des antennes des objets.

TAKE AWAY IoT

COMMUNIQUER

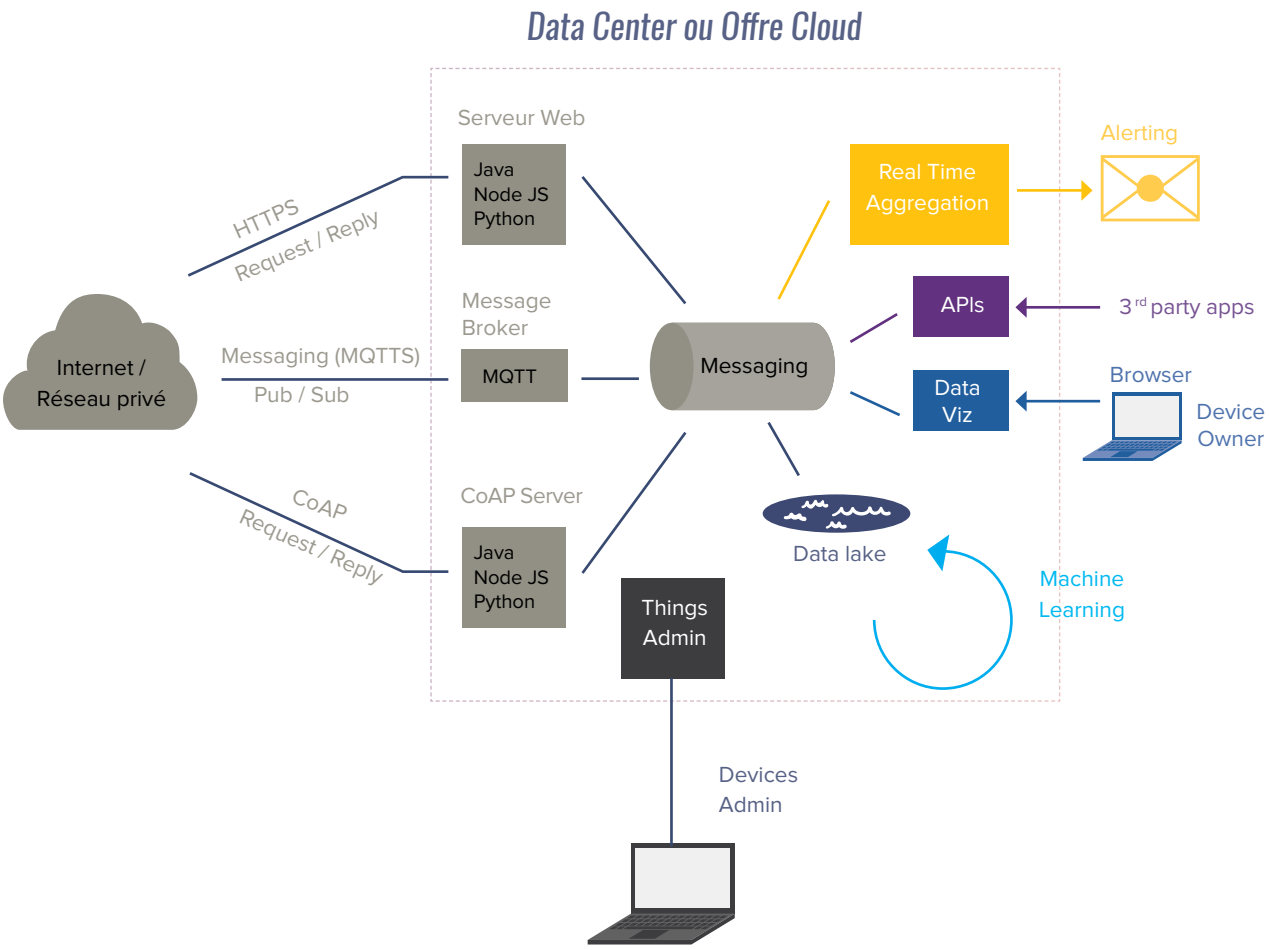
- Les objets communicants connaissent un développement exponentiel grâce à deux avancées majeures : la baisse continue des coûts de production des devices, et l'amélioration spectaculaires des moyens de communication disponibles.
- Les objets peuvent communiquer classiquement sur les réseaux existants : cellulaires et WiFi, mais ceux-ci font porter des contraintes fortes sur les durées de vie des batteries.
- De manières plus novatrice, on a vu l'apparition de réseaux dédiés à l'IoT, les LPWAN. Ceux-ci limitent les volumes de données transmis et la fréquence des échanges, mais permettent d'envisager des objets autonomes en énergie sur une période de plusieurs années. Les deux réseaux majeurs de cet écosystème sont Sigfox et LoRa.

TRAITER

Une fois les protocoles de télécommunication choisis,

il est temps de s'attaquer à l'architecture du Système d'Information sous-jacent. On entend souvent parler d'architecture IoT.

Or, il faut reconnaître qu'il n'y a rien de novateur dans ces architectures. Elles empruntent les meilleures pratiques des différentes innovations technologiques récentes : **BigData, NoSql, architectures de messaging, traitement temps réel, Cloud, etc.** Certes, les données collectées présentent des spécificités mais il serait abusif de parler de révolution du SI.



Architecture de traitement des données IoT

Des données particulières

Les données issues des objets connectés présentent, pour les architectes informatiques, des caractéristiques fortes :

- **Des données peu volumineuses mais très nombreuses.** Certains capteurs peuvent émettre plusieurs dizaines de valeurs toutes les minutes. La multiplication des capteurs sur le territoire national, voire à l'international, engendre des quantités de données qui se rapprochent de l'un des trois V du BigData[5]: le volume.
- **Des mesures de capteurs souvent organisées en séries temporelles.** Les objets connectés sont fréquemment pensés pour être suivis sur une longue période de temps (parfois plusieurs années).
- **Des formats de données très variés.** Il est toujours étonnant de voir comment les fabricants tentent systématiquement d'imposer leurs propres normes. On se retrouve ainsi avec presque autant de manières de représenter une température que de capteurs sur le marché. Là encore, nous tendons vers une autre caractéristique du BigData : la variété.

De plus, l'aspect temps réel de la plupart des objets et la nature des données qu'ils représentent nécessitent parfois une réaction rapide du système sous-jacent (prenons l'exemple du détecteur de fumée). Cette réactivité nous fait pencher pour une architecture orientée événements, qui permettra de corréler certaines constatations et de leur associer des chaînes de traitement idoines.

> *Les données issues des objets connectés présentent, pour les architectes informatiques, des caractéristiques fortes : des données peu volumineuses mais très nombreuses, des mesures de capteurs souvent organisées en séries temporelles et des formats de données très variés.*

5- Les 3V du Big Data : Volume, Vitesse et Variété.

Les protocoles de transport de la donnée

Au delà de la technologie choisie pour la communication, il est important de comprendre le rôle des protocoles de transport, qui vont utiliser les canaux physiques pour transférer la donnée sur Internet.

Parmi la multitude des protocoles existants, nous allons nous attacher à trois protocoles en particulier qui sont les plus utilisés dans le monde de l'IoT :

- **HTTP:** HyperText Transfer Protocol
- **MQTT:** Message Queuing Telemetry Transport
- **CoAP:** COnstrained Application Protocol

HTTP

Inutile de présenter HTTP dans les détails, il est la norme. Pourtant, dans le monde des objets connectés, il constitue un cas à part car très peu d'objets l'utilisent nativement. Il y a plusieurs raisons à celà. Tout d'abord, faire tourner un serveur HTTP requiert une certaine puissance de calcul, puissance d'autant plus importante si l'on souhaite sécuriser ce protocole via des certificats SSL. HTTP se voit donc quasiment disqualifié d'office. Néanmoins, on ne peut pas l'écarter. En effet, il reste souvent le vecteur choisi pour communiquer entre un fournisseur telco (e.g., Sigfox ou LoRa) et Internet, la plupart des opérateurs utilisant des callback HTTP entre leurs datacenters et la destination finale des données.

MQTT

MQTT est un protocole de messagerie publish-subscribe basé sur TCP/IP. MQTT est une extension du protocole MQ (Message Queuing) orienté spécialement vers les objets. Sa légèreté, la possibilité de le sécuriser (via MQTTS), **sa bidirectionnalité et les nombreux mécanismes de fallback et d'abonnement qu'il propose en font le premier choix** pour les objets communiquant directement sur Internet. Il constitue d'ailleurs la brique sous-jacente de tous les SDKs IoT du Cloud : AWS, Bluemix ou Azure. MQTT est un standard OASIS depuis novembre 2014.

CoAP

CoAP est un protocole applicatif qui tourne sur UDP. Il est fortement inspiré de REST et vise à contourner les limitations de HTTP. Il est **nettement plus léger, intègre nativement des mécanismes de sécurité et propose des fonctionnalités originales,** telles que la découverte automatique des noeuds appartenant au réseau.

Depuis juillet 2013, CoAP est un standard IETF. Il est encore peu exploité du fait de sa relative jeunesse. Il rencontre néanmoins un net succès parmi les industriels. Il pourrait d'ailleurs rapidement devenir une norme largement utilisée. Du point de vue du développeur, utiliser HTTP REST ou CoAP revient quasiment au même et cela facilite grandement son intégration.

Quel protocole pour mon objet ?

Comment choisir parmi tous ces protocoles ? Généralement ce choix sera dicté par votre manière de communiquer :

- Si vos objets dialoguent directement avec votre backend, vous avez tout intérêt à privilégier MQTTS en TCP ou CoAP en UDP.
- Si vos objets dialoguent avec le backend d'un fournisseur télécom, il est plus que probable que HTTPS vous soit imposé.

Architecture orientée événements

L'architecte IoT devra souvent mettre en place un système dit évènementiel : en fonction de la nature des données qu'il reçoit, plusieurs boucles de traitement peuvent être empruntées.

Ce type d'architecture repose massivement sur l'utilisation de mécanismes de publication / souscription, et les technologies de messaging. On peut alors réaliser des systèmes largement découplés qui peuvent croître facilement au fil du temps. À l'heure actuelle, **les architectures microservices[6] sont privilégiées**, permettant à un grand nombre de petits composants de réaliser des traitements riches sans pour autant sacrifier la flexibilité globale du système.

6- Le TechTrends dédié aux microservices est disponible en téléchargement pdf et epub sur xebia.fr

Les contraintes imposées par ce type de développement ne changent pas : le système doit être testé en profondeur, il est souhaitable de disposer de mécanismes de déploiement automatisé, et les mécanismes de surveillance du fonctionnement global sont un enjeu majeur. La question d'imposer un format pivot aux données divise encore les acteurs de l'écosystème. Néanmoins, les évolutions technologiques tendent à favoriser sa disparition (en particulier l'avènement de Big Data qui permet de traiter des données hétérogènes et faiblement structurées).

Bref, rien que du très connu mais discipline et rigueur restent indispensables.

Big Data et séries temporelles : un stockage à adapter aux usages

Vos données seront le coeur de votre système. À ce titre, elles doivent être stockées dans un réceptacle adapté. Les bases de données orientées séries temporelles tiennent ici la corde. En effet, **ces bases NoSql sont particulièrement adaptées pour stocker des couples valeur - date.** De plus, elles sont pensées pour réaliser des opérations complexes sur un grand nombre de tuples. Initialement dédiées au stockage des données de surveillance des machines du Cloud (CPU, RAM, etc.), ces dernières ont trouvé un regain d'intérêt en stockant des mesures en provenance des objets du quotidien (temperature, voltage, luminosité).

On retrouve ici les grands noms des bases de données orientées séries, à savoir InfluxDb, Graphite ou encore OpenTSDB. On citera aussi le projet open source de la startup brestoise (cocorico) Cityzen Data, Warp10, qui mérite toute notre attention du fait de ses concepts novateurs, en particulier la possibilité de travailler au niveau de la géolocalisation.

Des traitements comparables à ceux du Big Data

Outre ces bases de données qui permettent un usage immédiat de la donnée, il est fréquent de coupler ce datastore chaud avec un datastore froid dans lequel sera conservé une très grande profondeur de données à la fois dans le temps et dans la granularité. **Cette dimension d'historisation nous renvoie à l'architecture à la base de tous les projets Big Data : le DataLake.** Il devient alors naturel de mentionner Hadoop, Spark ainsi que la galaxie d'outils à la disposition des Data Engineers et des Data Scientists[7]. Ce dernier métier a alors un rôle tout particulier à jouer.

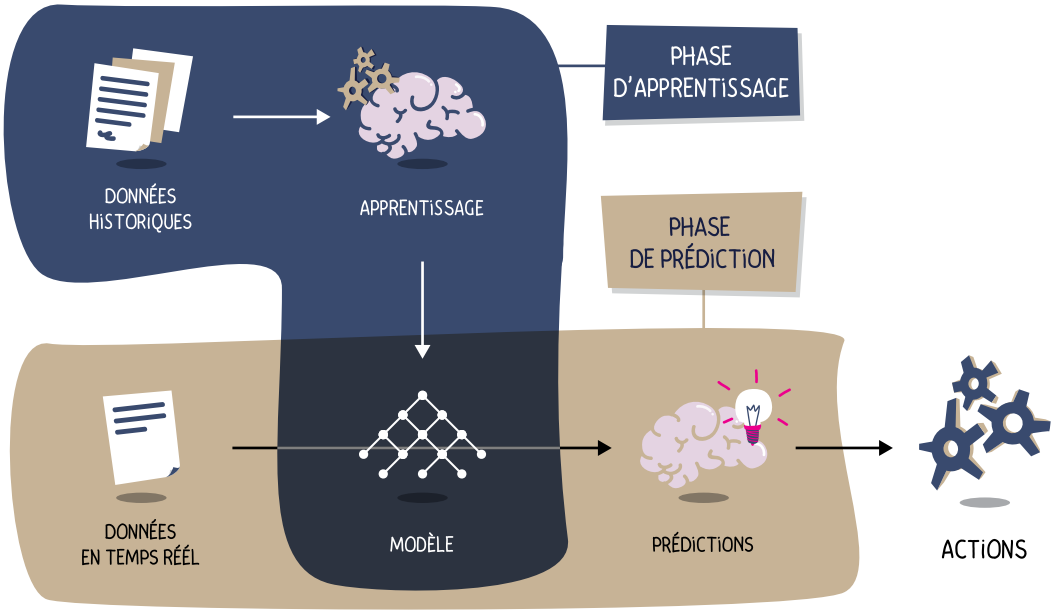
Le Machine Learning, un outil puissant

Maintenant que les données sont stockées, il est grand temps de les analyser et de les exploiter. La Data Science va permettre d'amener énormément de puissance pour ces traitements et de donner une valeur métier à ces données.

Rappelons qu'en Data Science, il est courant d'entraîner un modèle à partir d'un historique de données assez large, pour ensuite l'appliquer sur de nouvelles données, sur des plages plus ou moins longues en fonction du besoin. L'utilisation de cet historique va nous permettre d'avoir à notre disposition les principaux patterns présents dans la donnée, pour ensuite faire des prédictions pertinentes sur de nouvelles données.

> *La Data Science va permettre d'amener énormément de puissance pour ces traitements et de donner une valeur métier à ces données.*

7- Le TechTrends Data Lab est disponible en téléchargement pdf et epub sur xebia.fr



Étapes d'apprentissage d'un modèle de machine learning

Pour rendre cette théorie plus parlante, illustrons-la par un exemple.

Cas d'usage : Détection d'anomalies

Prenons le cas de la détection d'anomalies dans le fonctionnement de machines industrielles. Ces machines possèdent des centaines de capteurs, chacun remontant des informations diverses et variées sur leur état de fonctionnement et leur activité. L'objectif étant de détecter les éventuelles anomalies dans leur fonctionnement, nous allons nous baser sur l'historique des données recueillies pour que **le système apprenne des règles qui nous permettront de déceler une anomalie de fonctionnement sur les données des activités en cours** (et donc en temps réel).

Deux possibilités s'offrent à nous :

- L'historique des anomalies rencontrées est aussi sauvegardé avec les données. On peut alors se baser sur cet historique pour reconnaître les patterns inhérents à la présence d'anomalies, et savoir les reconnaître pour anticiper les prochaines au fil de l'eau.

- Si l'information sur les anomalies n'est pas recensée, on peut toujours changer notre fusil d'épaule et modéliser le comportement normal des machines à partir de l'historique des données récoltées. Les anomalies correspondent alors aux données pour lesquelles l'écart par rapport à ce comportement normal dépasse un certain seuil.

Dans tous les cas, un modèle de Machine Learning va pouvoir être entraîné sur des plages de temps longues. Une fois le modèle entraîné, il peut être exporté et utilisé en phase de prédiction, sur des plages de temps plus courtes ou même en temps réel.

L'utilisation d'algorithmes de Machine Learning n'a rien de spécifique à l'IoT. Les principes et méthodologies s'appliquent dans tous les domaines, et l'utilisation de séries temporelles est déjà commune dans beaucoup de cas d'applications. **Toute une panoplie d'algorithmes peut être utilisée selon les cas,** allant de la simple régression linéaire jusqu'au Deep Learning pour la reconnaissance d'images ou de sons.

Choisir d'avoir recours à la Data Science, c'est tout d'abord éviter d'avoir à gérer un format pivot. **La quantité de données collectées doit permettre de voir émerger rapidement des patterns techniques.** De ces patterns découlent la nature de la donnée. Bien que les fabricants d'objets gèrent chacun leur propre format, une température extérieure sera toujours représentée sous forme d'un nombre décimal compris entre -50 et +50 (degrés Celsius, si on se limite à la plaque européenne), alors qu'une luminosité se situera entre 0 et 50000 (lux). De plus, les variations journalières et saisonnières peuvent aussi permettre de rapprocher des données dont la représentation informatique diffère.

Enfin, la multiplication des capteurs, qu'ils soient environnementaux ou personnels, va permettre de nouvelles corrélations qui pourront renseigner de manière très précise sur les habitudes de chacun. Mais, comme nous le verrons, ces corrélations poseront rapidement des problèmes éthiques.

Exposition d'API

Après la collecte et le traitement vient le temps de la restitution et de la mise à disposition des données aux consommateurs. La restitution peut prendre de multiples formes, que ce soit des tableaux de bord dans une application Web, des widgets pour une application mobile voire même des données brutes pour un système décisionnel plus complexe.

Là encore, inutile de réinventer la roue. **L'exposition / consommation d'API est devenue un standard du marché et il serait déraisonnable de vouloir s'en passer.**

Cette manière d'exposer les données rendra visible l'Internet des Objets, que ce soit à travers de l'Open Data (comme par exemple la carte collaborative des possesseurs de thermomètres Netatmo[8]) ou bien de la monétisation de certaines données, comme ce qui peut exister à travers divers systèmes B2B[9].

Vers une architecture décentralisée

L'apparition des crypto-monnaies et des technologies de Blockchain décuplent la créativité de certains technophiles convaincus. En effet, pourquoi ne pas imaginer un monde où les objets effectuent des transactions simples de gré à gré (comme un frigo se réapprovisionnant de manière autonome pour les denrées de base).

Les technologies sont aujourd'hui disponibles mais, de notre point de vue, sans avoir atteint pour le moment un niveau de maturité suffisant. La débâcle DAO suffit pour s'en convaincre[10].

Pourtant, il ne faut pas condamner définitivement cette option. Une fois résolues les lourdes problématiques de sécurité, une telle décentralisation pourrait offrir un terrain de jeu particulièrement innovant.

8- https://weathermap.netatmo.com
9- https://www.farmobile.com/
10-http://www.lemondeinformatique.fr/actualites/lire-blockchain-l-attaque-contre-the-dao-conduit-ethereum-a-proposer-un-fork-65194.html

Les fonctions d'administration

Pour revenir à quelque chose de plus spécifique, il faut s'intéresser au cycle de vie des objets. En effet, l'idée est de lâcher dans la nature des systèmes informatiques miniaturisés, dont la durée de vie peut dépasser la dizaine d'années. Il est de notoriété publique que l'informatique n'est pas la championne de la stabilité dans le temps ...

Les fonctions d'administration et de pilotage des objets vont donc rapidement devenir le nerf de la guerre pour gérer la myriade d'objets en activité.

Encore une fois, aucun standard n'a pour le moment vraiment émergé et il existe plus de problèmes à résoudre que de solutions normées.

Les problématiques sont néanmoins globalement identifiées :

- **La surveillance des objets.** Fonctionnent-ils correctement ? Sont-ils aptes à remonter des données à la fréquence voulue ? Quel est l'état de leur batterie ? Ces questions peuvent être facilement résolues par un monitoring standard, encore faut-il l'avoir prévu.
- **Le déploiement.** Si les devices ne sont pas géolocalisés, comment puis-je associer un objet déployé dans un lieu donné à un identifiant unique ? En cas de dysfonctionnement d'un objet, quelle sera la politique d'intervention (sur site, retour atelier, abandon et remplacement, etc.) ?
- **La mise à jour.** Les normes informatiques évolueront à coup sûr, comment puis-je mettre à jour le firmware de mon objet (en particulier s'il n'est doté que d'une communication unidirectionnelle) ? Comment tester qu'une mise à jour logicielle ne fera pas disparaître l'objet de ma grille ?
- **La détection de fraude.** Si mon objet est physiquement captif d'une entitée malveillante, comment puis-je le savoir ? Suis-je en mesure de l'isoler de mon réseau ? Suis-je en capacité de l'empêcher d'envoyer des données erronées sur mon système ? Le logiciel déployé sur cet objet permet-il de rétro-concevoir mon système dans sa globalité, et donc de l'exposer à une attaque plus importante ?

Des solutions existent pour chacune de ces questions mais il est actuellement difficile d'établir des normes, chaque objet et chaque fabricant traçant sa propre route.

Tout faire soi-même ou prendre une offre sur étagère ?

Les derniers trimestres ont vu la multiplication des offres dédiées à l'IoT, qu'elles proviennent des opérateurs télécom tels Bouygues Telecom ou Orange, d'acteurs Cloud comme AWS ou OVH, ou bien d'acteurs historiques du Machine-2-Machine (M2M) comme Bosch ou Ericsson.

Comme nous l'avons vu, l'architecture IoT n'a, en fait, rien de réellement novateur. **L'arbitrage entre un système fait maison et une offre clés en main se pose donc dans les mêmes termes que n'importe quel projet informatique.**

Gageons néanmoins que les acteurs du Cloud ont sur ce sujet plusieurs coups d'avance. En effet, ils offrent pour la plupart des caractéristiques permettant d'accélérer drastiquement les développements d'une plateforme IoT :

- **L'élasticité,** permettant de s'adapter rapidement à la croissance d'une flotte d'objets.
- **De multiples services en PaaS,** tant au niveau des bases de données orientées séries que des bus Pub/Sub ou encore de la gestion d'APIs.
- **Des offres payables** à la consommation simplifiant la construction rapide de PoC à moindre coûts.
- La possibilité de **combiner du PaaS et du IaaS** pour s'adapter à tous les besoins.

Les offres évoluant bien plus vite que notre rythme de publication, nous ne nous lancerons pas dans un comparatif exhaustif. Mais le nombre d'acteurs et la richesse des propositions croissant de manière exponentielle, nous ne doutons pas que chaque projet pourra trouver un package avantageux n'obligeant pas à recourir à des développements spécifiques on-premises.

TAKE AWAY IoT

TRAITER

- L'IoT repose majoritairement sur des briques du SI déjà connues depuis quelques années : HTTP et MQTT, les architectures réactives, BigData et la DataScience, les API, le Cloud.
- Même si les briques de base sont connues, leur agencement et leur pleine exploitation, peuvent rapidement devenir une gageure pour un non spécialiste.
- Concevoir un backend IoT implique de maîtriser quelques spécificités : il est nécessaire de savoir gérer le cycle de vie des objets et donc de disposer d'un minimum de reporting et de fonctions d'administration.

SÉCURISER

Nous l'avons vu, l'Internet des Objets a vocation à s'immiscer partout :

à domicile, pendant nos activités quotidiennes, à l'intérieur des bâtiments industriels et des bureaux, dans les transports, etc. Le corollaire est que, plus il y aura de données et d'usages, plus ce nouveau champ d'application deviendra la cible des pirates de tous bords.

Une nécessité vitale

Il est donc crucial, dès maintenant, de penser la sécurité en amont afin de diminuer les surfaces d'attaques et de décourager les hackers potentiels. Le principe de base de toute mise en oeuvre sécurisée est de **penser la sécurité dès le design[11] du système**. Plutôt que de dissimuler les mécanismes sécuritaires qui régissent le software, il est préférable de les exposer afin que les utilisateurs puissent influer directement sur le logiciel avant qu'il ne soit trop tard. C'est d'ailleurs le parti pris de l'Alliance LoRa, qui décrit ses protocoles de sécurité dans ses spécifications.

De plus, tous les cas ne sont pas comparables. Lorsque la presse papier française a perdu une partie de sa base d'abonnement, l'impact sur les utilisateurs était relativement bien circonscrit. En 2016, Sony s'est fait dérober des numéros de carte de crédit dans le PlayStation Store. L'impact est ici plus incertain, cela pouvant entraîner des dommages directs sur le compte en banque des joueurs. Heureusement pour Sony, ses abonnés sont visiblement plus intéressés par les contenus que par la sécurité de leurs données bancaires, et le Store n'a curieusement pas souffert de cet épisode.

Imaginons, dans un futur pas si éloigné, que les équipements médicaux soient massivement connectés : en cas de fuite, ce ne sont plus de simples comptes bancaires qui seront menacés, mais la vie quotidienne de milliers d'individus explicitement nommés.

L'IoT, de par ses concepts, implique de faire communiquer un grand nombre d'objets, tout en facilitant leur accès et leur utilisation par le grand public. On parle ici de millions d'objets répartis sur une vaste zone géographique : **les zones d'attaques sont donc multipliées**. Ainsi, chaque device devient une porte d'entrée vers des Systèmes d'Informations plus larges, mais peut aussi devenir un relai pour des attaques plus classiques. Le 21 octobre 2016, le service de gestion de DNS dyn.com[12] a subi une attaque de type déni de service massive, orchestrée via plusieurs millions de dispositifs, tels que caméras IP ou box internet, infectés par le virus Mirai[13]. L'étendue de l'attaque (on parle de plus de 10 millions d'IP) et sa complexité (les IP sources étaient répandues sur l'ensemble de la planète, et provenaient de systèmes variés) donnent une idée de ce que les cybercriminels seront en mesure de faire dans quelques années si l'on n'y prend pas garde.

11- Secure by design : https://en.wikipedia.org/wiki/Secure_by_design

12- http://dyn.com/blog/dyn-statement-on-10212016-ddos-attack/

13- https://en.wikipedia.org/wiki/Mirai_(malware)

Comme pour toute dette technologique, il faudra un jour payer le retard accumulé sur le plan de la sécurité. Dans un marché d'innovation toujours plus rapide, les projets les plus pérennes seront également les plus responsables sur la sécurisation de leurs produits.

Le monde de la sécurité possède aussi ses formations.[14]

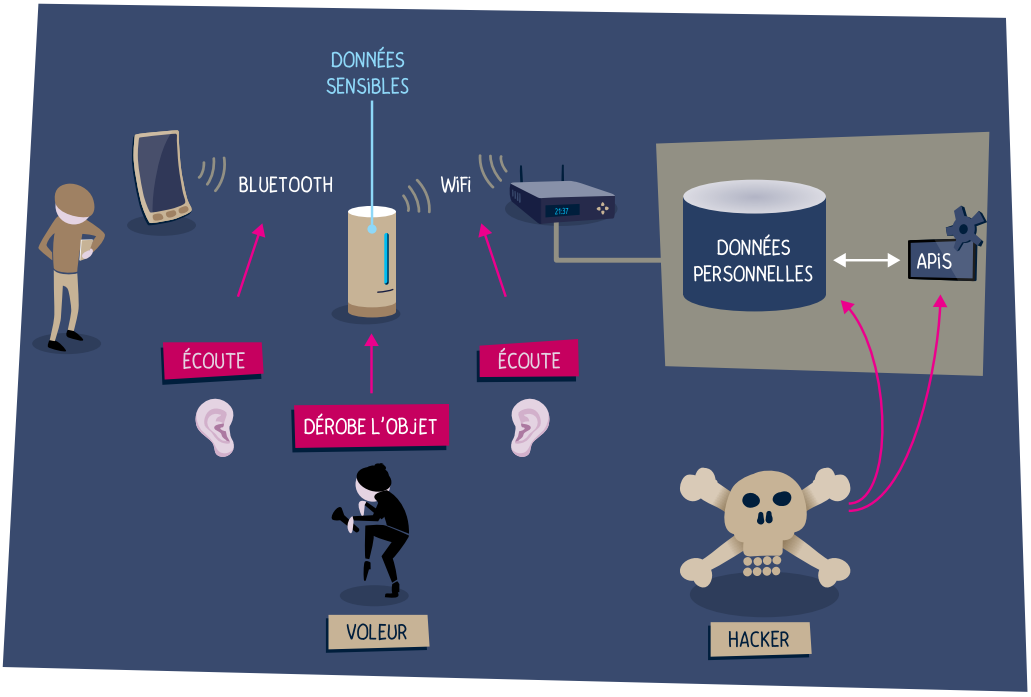
Les axes de sécurisation que nous allons aborder ne sont pas nouveaux. **Aucun système ne peut se dire 100% sécurisé.** Le challenge ici est bien de **rendre le coût de piratage plus élevé que les bénéfices potentiels.** Si les données que vous manipulez ne concernent que des informations triviales, vous ne faites pas courir de grands risques à vos utilisateurs. En revanche, dès que vos objets permettent de rentrer dans la vie privée des gens ou d'agir sur le monde physique, comme avec des caméras ou des interrupteurs connectés, tout devient plus délicat.

La première étape de votre démarche de sécurité commence par **évaluer le Modèle de Menace**[15] : il faut définir, du point de vue de l'attaquant, quelles informations seraient susceptibles d'être intéressantes et quels vecteurs d'attaque permettraient d'y accéder. Une méthodologie émergente pour analyser le Modèle de Menace est le VAST (Visual, Agile, and Simple Threat modeling) : elle vise à intégrer le processus de Modélisation des Menaces au sein des équipes DevOps, dans un souci de co-construction itérative.

De manière générale, de la même façon que l'agilité a transformé l'équipe de développement et a entraîné la naissance de DevOps, il faut maintenant intégrer la dimension sécurité dans l'équipe pour aller vers des équipes pluri-disciplinaires DevSecOps.

14- http://www.securitytube-training.com/online-courses/offensive-internet-of-things-exploitation/index.html
15- https://en.wikipedia.org/wiki/Threat_model

Axes de travail



Les zones d'attaque pour un hacker dans l'IoT

Données en transit

Les objets connectés communiquent sur des canaux communs, Bluetooth ou WiFi, à l'exception notable des réseaux LPWAN. **La zone d'attaque la plus évidente consiste donc à intercepter la communication entre un objet et son backend.** Afin de mettre leurs devices à la disposition du plus grand nombre, les fabricants simplifient grandement la connexion du device avec Internet. Le paramétrage doit, en priorité, demander un effort minimal de la part de l'utilisateur.

Bluetooth et WiFi sont des technologies éprouvées et répandues, mais aussi largement maîtrisées et documentées par les hackers. Il n'est donc pas difficile de détourner leur usage si elles sont mal protégées.

Le point de rencontre des objets connectés étant souvent le smartphone de l'utilisateur, une première mesure pourrait être de n'autoriser le paramétrage qu'aux personnes physiquement proches de l'objet. Un appairage Bluetooth peut par exemple exiger l'appui sur un bouton physique à l'image des thermomètres Netatmo ou des balances Withings.

Une fois le téléphone et l'objet appairés, il est bien sûr nécessaire de chiffrer les flux. Aucun autre device Bluetooth dans la zone ne doit pouvoir récupérer les informations échangées. Avec l'apparition du standard Bluetooth LE, le chiffrement AES devient la norme, ce qui permet un niveau correct de sûreté de communication.

La communication entre l'application présente sur le smartphone et une plateforme centrale doit également être l'objet d'attention. **Toute communication se doit d'être chiffrée et authentifiée par certificat.** Ainsi, il est judicieux de mettre en place des mécanismes de Public Key Pinning, qui n'autorisent vos applications à n'accepter que des certificats SSL définis préalablement ou, encore, en tant qu'éditeur, de créer votre propre autorité de certification.

L'utilisation des réseaux LPWAN apporte un autre type de réponse. Le réseau est administré de bout en bout, l'objet connecté bénéficie d'un paramétrage en usine pour se connecter directement et sans heurts, et les couches de chiffrement utilisées y sont nombreuses. Les spécifications LoRa en sont d'ailleurs un bon exemple. Un petit défaut néanmoins : à l'heure actuelle, il est assez aisé de se procurer une antenne LoRa pour devenir une gateway publique. En interceptant la totalité du trafic (et donc en ayant suffisamment de données), il devient envisageable de casser les protocoles d'encryption. Rappelez vous qu'aucun chiffrement n'est inviolable, et le renouvellement périodique des clés de chiffrement sur l'ensemble de la chaîne applicative n'est donc pas une option.

Données au repos

Après le transit, votre donnée sera enregistrée afin d'être consultée ultérieurement. De ce côté là, rien de neuf. Stocker des données de mesure n'est pas plus compliqué que de les enregistrer dans un DataLake. Néanmoins, rappelons à tout hasard que **la sécurisation d'un DataLake est souvent la partie la plus ardue d'un projet data,** et que la gestion fine des données et des accès en est souvent le parent pauvre. Il en va de même pour l'IoT : la sécurisation des silos de données est un enjeu clé et reste une tâche non négligeable au sein du projet.

La gestion des données au repos présente plusieurs complexités :

- Le contrôle d'accès en est la partie émergée. Une fois les données de mesure collectées, il faudra filtrer qui peut (ou non) les voir. Certaines informations peuvent relever directement de la vie privée, et sont soumises à des lois souvent complexes. La transposition de ces lois est souvent un réel casse-tête technique.

- L'audit et la traçabilité seront bien sûr des outils de contrôle demandés par toutes les instances légales, comme la CNIL, mais aussi par les instances de contrôle internes. Toute intervention sur les données, que ce soit en lecture ou en modification, doit être tracée, et auditable, de manière simple et efficace. La multiplication des données, des sources et des intervenants, peut faire que les données d'audit représentent un volume plus conséquent que les données réelles.

Les sujets sont nombreux, et demandent chacun une expertise pointue et un temps de réalisation non négligeable.

Authentification du porteur

Déployer massivement et sur des zones étendues des devices amène à s'exposer tôt ou tard au vol de ces derniers. Mais que se passe-t-il quand une personne mal intentionnée met la main sur un dispositif ? Devient-il un risque potentiel pour mon organisation ou bien ne devrais-je gérer que l'impact financier du vol ? Dans tous les cas, il faut s'y préparer.

Tout d'abord, les informations sensibles résidant sur le device doivent être inattaquables. **Toute information stockée doit être chiffrée**, idéalement de manière asymétrique. Ensuite, l'empreinte de ces données doit être minimale. **Ne stockez que le strict nécessaire.** Au besoin, mettez à disposition des canaux d'information dédiés pour obtenir les paramètres non stockés. Par exemple, certaines informations critiques de configuration peuvent être chargées, au démarrage, depuis un point sécurisé et ne résider qu'en mémoire vive. Ainsi, vous rendrez leur extraction bien plus complexe.

Il existe enfin des mécanismes de protection du firmware même de l'objet le rendant inutilisable en cas d'intrusion. Néanmoins, ce type de protection ne pourrait pas s'appliquer à tous les cas d'usage. Sans même s'attaquer aux données stockées, un objet échappant à votre contrôle peut devenir dangereux, en détournant son usage nominal (refroidir artificiellement un capteur de température afin de déclencher le thermostat par exemple).

Auteurs : Martin Untersinger et Thibaut Soulcié. Publié dans "LA REVUE DESSINÉE" #14[16]*, automne 2016*

À tout moment, vous devez être en mesure de gérer votre flotte d'objets (leur localisation, leur niveau de batterie, leur étalonnage, leurs paramètres vitaux) et surtout de les empêcher d'interagir avec votre backend. Les acteurs du Cloud et les anciens du M2M proposent à ce titre de provisionner les objets avec des certificats qu'il est possible de révoquer à distance. Ainsi, un objet perdu continuera à communiquer sur vos canaux, mais ne sera plus reconnu comme appartenant à votre organisation.

Enfin, **la Data Science peut aussi nous aider à écarter les comportements suspects.** Si un objet se comporte de manière anormale, il est possible de détecter un pattern correspondant à un détournement de son usage : envoi de données plus fréquent, géolocalisation aléatoire, etc.

Un métier à part entière

Les métiers de la sécurité sont souvent méconnus, même au sein de la sphère IT. Sécuriser ses applications, ses communications et son stockage a trop longtemps été un poste négligé. Si aujourd'hui on peut comprendre que les équipes aient mis l'accent sur l'acquisition de marché et les fonctionnalités visibles, nous parions que, à l'ère de l'hyper-connexion, la sécurité et la confidentialité seront les avantages concurrentiels de demain.

16- http://www.larevuedessinee.fr/Numero-14

Les méchants gagnent parce que les gentils n'imaginent pas jusqu'où ils peuvent aller. La veille technologique nécessaire pour se tenir à jour sur les attaques potentielles et les techniques de remédiation est un métier à plein temps. En réalité, il s'agit même de plusieurs métiers à plein temps ...

XXX-Sec

Comme pour les compétences développement ou système, il est souvent complexe de ranger les informaticiens dans des cases bien bornées. On voit néanmoins se dégager quatre grands rôles dans le monde de la sécurité :

• Information Security (InfoSec)

La sécurité n'est pas qu'une affaire de technologie. Si les utilisateurs ne sont pas sensibilisés aux risques et aux attitudes à avoir, ils resteront à la merci du premier phishing ou des débutants en social-engineering. Le volet InfoSec est dédié à la sensibilisation et à l'éducation de tous les utilisateurs, de vos clients finaux à vos opérateurs. Intégrer un expert InfoSec à vos équipes de communication peut être un avantage considérable.

• IT Security (ITSec)

Il est courant dans les architectures IoT de trouver une forme quelconque de plateforme centrale, un point de rendez-vous des données et une ou plusieurs applications Web pour y accéder. L'ITSec va se concentrer sur la sécurisation de ces systèmes d'informations.

• Network Security (NetSec)

Firewall, chiffrement, encapsulation de protocoles, tout ce qui passe au travers d'équipements réseaux, fermés ou ouverts, relève du domaine du Network Security.

• Application Security (AppSec)

La validation des données saisies, la solidité face à des tentatives d'escalade de privilège, de flooding ou d'injections en tout genre, sont du ressort du praticien en Application Security.

Blue Team / Red Team

Avoir des spécialistes en sécurité en interne et les faire travailler sur vos objets connectés ne suffit pas en soi. Personne n'est infaillible. Il faut donc faire valider le travail de vos internes, qui constituent la Blue Team (les gentils) par des équipes externes qui rempliront le rôle d'attaquants : la Red Team (les méchants).

Blue et Red Team peuvent faire partie de votre organisation, mais ils n'ont pas la même mission. Ne vous y trompez pas, votre Red Team sera bien plus pertinente si vous partagez toutes les informations techniques de votre plateforme avec eux. Ils sont là pour trouver les failles du système. Ils trouveront bien plus vite des failles plus pertinentes avec un maximum d'information.

I am the cavalry, and so are you

Les professionnels du monde de la sécurité ont compris depuis un moment les enjeux que nous avons exposés ici. Les conférences comme DefCon[17] ou BlackHat[18], pour ne citer que les plus connues, regorgent de sujets sur les failles de l'industrie des objets connectés. La sécurisation est l'exception. Pour des professionnels de la sécurité, voir des industriels déployer des flottes d'objets impactant directement leur quotidien et faisant rentrer dans leur vie physique les mêmes risques qu'ils affrontent dans le monde numérique est une chose insupportable.

Le mouvement "I am the cavalry"[19] a été lancé en 2013 à Las Vegas, lors de la plus grosse conférence sécurité au monde, DefCon. Plutôt que de continuer à tourner l'IoT en dérision, des professionnels de toutes les spécialités se sont donné pour mission de faire partie de cet élan d'innovation pour le faire aller dans le bon sens. Certains ont même changé de travail pour se faire embaucher chez des industriels de l'IoT médical, automobile et infrastructures publiques, afin de changer les choses de l'intérieur.

Cette croisade contre le laisser-aller de la plupart des initiatives est louable et nous vous encourageons à vous rapprocher de ces acteurs pour obtenir conseils, ressources et contacts pour bien démarrer.

17- https://www.defcon.org/
18- http://www.blackhat.com/
19- https://www.iamthecavalry.org/about/history/

IoT et éthique

Nous ne saurions conclure ce paragraphe sur la sécurité sans y joindre une note éthique. Les techniques et les technologies mises à la disposition des industriels permettent de réaliser de grandes choses. Mais parfois, leur usage primaire s'accompagne d'usages secondaires et mercantiles qui peuvent rapidement influer sur la vie des citoyens.

Il est du devoir et du ressort de tous, producteurs de service comme consommateurs, de s'engager dans une voie éthique, qui permet aux citoyens de rester libres malgré leurs usages. Même si certaines commissions sont là pour garantir nos droits et nous protéger, les usages vont parfois plus vite que les lois qui les accompagnent.

Le législateur a par le passé montré ses limites et son insuffisance. Dans une économie globalisée, les lois ne peuvent être transfrontalières et universelles et ont démontré leurs failles (Google, Uber). L'IoT ne dérogera pas à ces dysfonctionnements.

Certains utilisent le terme de sousveillance[20] pour désigner l'activité de documenter la surveillance que peuvent opérer ces objets sur notre vie quotidienne dans tous les domaines : santé, localisation, habitudes, appétences par exemple.

Afin d'assurer cette sousveillance de l'Internet des Objets, il faudra s'appuyer sur quatre leviers :

- **Les technologies** pour assurer encryption et sécurité d'accès aux données privées.
- **Les lois pour gérer la protection de la vie privée,** le droit à l'oubli, l'encadrement de la collecte des données et la responsabilité pénale des acteurs de l'écosystème.
- Le marché pour **adopter des codes de conduite éthiques** par la pression concurrentielle.
- **Les groupes de réflexion sociétaux et associations de consommateurs** pour influencer la pensée collective, éduquer les utilisateurs et influencer l'émergence de nouvelles réglementations, car il en faudra.

L'agence de sécurité européenne (ENISA) s'est d'ailleurs saisie du sujet, et a rendu un rapport passionnant[21] sur les "Smart Homes" encourageant chaque industriel, mais aussi chaque citoyen, à se pencher avec énormément de sérieux sur ces problématiques.

20- https://fr.wikipedia.org/wiki/Sousveillance

21- https://www.enisa.europa.eu/publications/security-resilience-good-practices

TAKE AWAY IoT

SÉCURISER

- La sécurité des données de vos utilisateurs n'est pas une option. Concevez des applications "Secure by Design".
- Les surfaces d'attaque dans l'IoT sont nombreuses : pendant le transit des données, lors de leur stockage final, voire même lors de leur accès, par le biais d'une usurpation d'identité.
- La sécurité est un métier à part entière, et inclure des spécialistes de la question à vos équipes de développement est une façon idéale de sécuriser vos projets dès leur démarrage.
- Au delà des aspects liés à la sécurité, l'éthique des entreprises gravitant dans le monde de l'IoT, et l'utilisation de vos données privées, posent de vraies questions philosophiques sur l'avenir de nos sociétés. Nous devons tous être acteurs des orientations qui seront prises dans les années à venir.

CONCLUSION

L'internet des objets est un monde en pleine émergence et rempli de promesses.

Les avancées techniques sont nombreuses, au niveau du hardware, des réseaux de télécommunications et du logiciel.

Les principales avancées se situent dans **l'émergence de réseaux radio dédiés spécifiquement à l'émission / réception de nombreux paquets de données, d'une taille limitée.** Mais l'existence de ces standards naissants n'exclut absolument pas l'utilisation de canaux plus classiques, comme les réseaux cellulaires ou le WiFi.

Du côté de l'architecture logicielle, par contre, peu de nouveautés. **Récupérer, traiter et restituer la donnée reposent sur des bases connues :** Big Data, Cloud, Machine Learning, architectures évènementielles et outils de messaging. Certaines spécificités viennent parfois se greffer, mais on reste dans un domaine relativement bien balisé.

À l'inverse, on pourra s'étonner des **faibles percées sur le plan de la sécurité et de la confidentialité des données personnelles.** Sous-estimés ou bien volontairement délaissés, ces aspects sont autant de portes d'entrées pour de futurs scandales. Avec des chaînes de traitement de plus en plus sophistiquées, maintenir la sécurité de bout en bout devient complexe. Enfin, les aspects légaux et éthiques sont eux aussi à la traîne. Des associations de défense des droits des citoyens comme "La Quadrature du Net"[22] sont souvent démunies face aux appétits des grands industriels de l'information.

L'enjeu est d'autant plus important que les nouveaux usages et services offerts par l'IoT s'appliquent à un nombre toujours croissant des aspects de notre existence : de la simplification de la vie quotidienne à la réduction de l'impact énergétique, d'une infrastructure routière plus efficace et, donc, plus sûre, jusqu'à l'émergence d'une nouvelle medicine. Une chose est certaine : bien utilisés, les objets connectés pourront révolutionner bien des choses.

22- https://www.laquadrature.net/fr

TAKE AWAY IoT

01 COMMUNIQUER

- Les objets communicants connaissent un développement exponentiel grâce à deux avancées majeures : la baisse continue des coûts de production des devices, et l'amélioration spectaculaires des moyens de communications disponibles.
- Les objets peuvent communiquer classiquement sur les réseaux existants : cellulaires et WiFi, mais ceux-ci font porter des contraintes fortes sur les durées de vie des batteries.
- De manières plus novatrice, on a vu l'apparition de réseaux dédiés à l'IoT, les LPWAN. Ceux-ci limitent les volumes de données transmises et la fréquence des échanges, mais permettent d'envisager des objets autonomes en énergie sur une période de plusieurs années. Les deux réseaux majeurs de cet écosystème sont Sigfox et LoRa.

TRAITER

- L'IoT repose majoritairement sur des briques du SI déjà connues depuis quelques années : HTTP et MQTT, les architectures réactives, BigData et la DataScience, les API, le Cloud.
- Même si les briques de base sont connues, leur agencement et leur pleine exploitation, peuvent rapidement devenir une gageure pour un non spécialiste.
- Concevoir un backend IoT implique de maîtriser quelques spécificités : il est nécessaire de savoir gérer le cycle de vie des objets et donc de disposer d'un minimum de reporting et de fonctions d'administration.

SÉCURISER

- La sécurité des données de vos utilisateurs n'est pas une option. Concevez des applications "Secure by Design".
- Les surfaces d'attaque dans l'IoT sont nombreuses : pendant le transit des données, lors de leur stockage final, voire même lors de leur accès, par le biais d'une usurpation d'identité.
- La sécurité est un métier à part entière, et inclure des spécialistes de la question à vos équipes de développement est une façon idéale de sécuriser vos projets dès leur démarrage.
- Au delà des aspects liés à la sécurité, l'éthique des entreprises gravitant dans le monde de l'IoT, et l'utilisation de vos données privées, posent de vrais questions philosophiques sur l'avenir de nos sociétés. Nous devons tous être acteurs des orientations qui seront prises dans les années à venir.


TAKE AWAY
IoT
TRENDS
TAKE AWAY

Merci à

Simone Civetta, Qian Jin, Christophe Heubès,
Séven Le Mesle, Anne Beauchart
et François Laurain.

À lire et à relire

Les précédents numéros des TechTrends sont disponibles en téléchargement (pdf et epub) sur xebia.fr.

Si vous souhaitez recevoir une version papier, nous vous invitons à envoyer un mail à marketing@xebia.fr

LES AUTEURS

Pablo
Lopez

Aurélien
Maury

Damien
Baron

Yoann
Benoit

Sameh
Ben Fredj

Luc
Legardeur

IoT-ee

IoT DONE RIGHT

IoT-ee
156 bd Haussmann, 75008 Paris
+33 (0)1 53 89 99 99
contact@iot-ee.com

Toutes les informations sur :
iot-ee.com
iot-ee.com/blog