République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Faculté des Sciences Exactes et des sciences de la Nature et de la Vie

Département de Mathématique et Informatique

MEMOIRE DE MASTER

Domaine *: Mathématique et Informatique*

Filière *: Informatique*

Option *: Systèmes d'Information*

Thème :

La Sécurité Des Données Dans L'Internet Des Objets (IoT)

Présenté par:

Tebib Samia et Soualah Abdelghani

Devant le jury:

Zammar Ammar	***MAA***	***Université de Tébessa***	***Encadreur***
Benjanna A/Hakim	***MCA***	***Université de Tébessa***	***Président***
Bouregaa Salima	***MCB***	***Université de Tébessa***	***examinateur***

Date de soutenance: ../05/2017

Remercîments

A l'issue de ce travail, nous voulons d'abord remercier Allah de nous guider et donner la force, le courage et la patience pour tirer le meilleur parti de ce travail: MERCI "Dieu"

Nous tenons à exprimer nos profondes gratitudes et sincères remerciements à nos promoteurs Mr. Zemmar Ammar, pour avoir accepté diriger ce travail et pour leurs aides et leurs orientations.

Chaleureux remerciements vont à tous nos enseignants au cours de notre formation au fil des années.

Ce travail n'aurait pas été possible sans le soutien affectueux de plusieurs personnes.

Nous trouvons submergés en leur offrant tous nos remerciements à dédier ce travail pour eux. Par ailleurs, nous tenons à remercier vivement les membres de jury qui ont fait l'honneur d'accepter de participer à notre soutenance de ce mémoire.

Soualah Abdelghani *Tebib Samia*

ملخص

إن إنترنت الأشياء يترجم حاليا إلى زيادة في عدد الأجسام المتصلة، أي الأجهزة التي لها هويتها الخاصة، والحوسبة، والاتصال أكثر وأكثر تطورا: الهواتف والساعات والأجهزة المنزلية، وما إلى ذلك. وهذه الكائنات تشغل عددا متزايدا من أجهزة الاستشعار والمحركات مما يسمح لهم لقياس البيئة والعمل على ذلك، مما يجعل هناك صلة بين العالم المادي والعالم الافتراضي. وعلى وجه التحديد، يطرح إنترنت الأشياء عدة مشاكل، لا سيما بسبب حجمه الكبير جدا وطبيعة ديناميكية وعدم تجانس البيانات والنظم التي تجعله (قوة قوية / منخفضة، ثابتة / متنقلة، إمدادات الطاقة، وما إلى ذلك.)

وتتطلب هذه الميزات أدوات وأساليب مناسبة لتحقيق تطبيقات قادرة على استخراج معلومات مفيدة من مصادر البيانات المتعددة المتاحة والتفاعل مع البيئة، عن طريق المحركات، وكذلك مع المستخدمين، عن طريق واجهات مخصصة.

وقد شجع انتشار الشبكات المتنقلة، وأجهزة الاستشعار، على تطوير مفاهيم حاسوبية مستقلة مع طائفة واسعة من التطبيقات. ومع ذلك، فإن الضعف الكامن في هذه الشبكات المستقلة يدخل تحديات أمنية جديدة، مثل الهجمات الداخلية من قبل الكيانات الخبيثة. وكثير من هذه الهجمات يصعب الكشف عنها وإحباطها بسبب سلوكها المتقلب لسلوك العمليات المشروعة في النظم المتفاعلة.

وثمة قيد آخر يتمثل في الحد من موارد بعض الشبكات القائمة بذاتها (شبكات الاستشعار اللاسلكية، والشبكات المتنقلة المخصصة) على متانة هذه العمليات، التي تشمل كلا من الأخطاء والأمان.

والهدف من هذه المذكرة هو دراسة الجانب الأمني في إنترنت الأشياء واقتراح حل فعال لتأمين البيانات من خلال آليات فعالة

Résumé

L'Internet des objets (ou IdO) se traduit à l'heure actuelle par l'accroissement du nombre d'objets connectés, c'est-à-dire d'appareils possédant une identité propre et des capacités de calcul et de communication de plus en plus sophistiquées : téléphones, montres, appareils ménagers, etc. Ces objets embarquent un nombre grandissant de capteurs et d'actionneurs leur permettant de mesurer l'environnement et d'agir sur celui-ci, faisant ainsi le lien entre le monde physique et le monde virtuel. Spécifiquement, l'Internet des objets pose plusieurs problèmes, notamment du fait de sa très grande échelle, de sa nature dynamique et de l'hétérogénéité des données et des systèmes qui le composent (appareils puissants/peu puissants, fixes/mobiles, batteries/alimentations continues, etc.).

Ces caractéristiques nécessitent des outils et des méthodes idoines pour la réalisation d'applications capables d'extraire des informations utiles depuis les nombreuses sources de données disponibles et d'interagir aussi bien avec l'environnement, au moyen des actionneurs, qu'avec les utilisateurs, au moyen d'interfaces dédiées.

La prolifération des réseaux ad hoc mobiles, pair-à-pair et de capteurs ont encouragé le développement des concepts d'une informatique autonome avec potentiellement un large éventail d'applications. Or, la vulnérabilité inhérente de ces réseaux autonomes introduit de nouveaux challenges de sécurité, telles que des attaques internes menées par des entités malveillantes. Plusieurs de ces attaques sont difficiles à détecter et à contrarier en raison de leur comportement asymptotique au comportement de processus légitimes des systèmes en interaction. Par ailleurs, la limitation des ressources de certains réseaux autonomes (réseaux de capteurs sans fil, réseaux mobiles ad hoc) constitue un autre grand challenge pour leur robustesse qui englobe à la fois la tolérance aux défaillances et la sécurité.

L'objectif de notre mémoire est d'étudier l'aspect sécurité dans l'IoT et propose une solution effective pour sécuriser des données à travers des mécanismes efficaces.

Abstract

The Internet of things (IoT) is currently translated into an increase in the number of connected objects, i.e. devices having their own identity, computing, and communication more and more sophisticated: telephones, watches, appliances, etc. These objects embark a growing number of sensors and actuators allowing them to measure the environment and act on it, thus making the link between the physical world and the virtual world. Specifically, the Internet of objects poses several problems, in particular because of its very large scale, its dynamic nature and the heterogeneity of the data and the systems that make it up (powerful / low power, fixed / mobile, Power supplies, etc.).

These features require tools and methods suitable for the realization of applications capable of extracting useful information from the numerous data sources available and interacting with the environment, by means of the actuators, as well as with the users, By means of dedicated interfaces. The proliferation of ad hoc mobile, peer-to-peer and sensor networks have encouraged the development of autonomous computing concepts with potentially a wide range of applications. However, the inherent vulnerability of these autonomous networks introduces new security challenges, such as internal attacks by malicious entities. Many of these attacks are difficult to detect and thwart due to their asymptotic behavior to the behavior of legitimate processes in interacting systems.

Another limitation is the limitation of the resources of some stand-alone networks (wireless sensor networks, ad hoc mobile networks) to their robustness, which encompasses both fault tolerance and security.

The objective of this thesis is to study the security aspect in IoT and propose an effective solution to secure data through efficient mechanisms..

Glossaire

GPS: Global Positioning System

CSP: Contrat de Sécurisation Professionnelle

SAP: Systems, Applications and Products for data processing

LINQ: Language Integrated Query

HPC: High performance computing

HDFS: Hadoop Distributed File System

CSA : Conseil Supérieur de l'Audiovisuel

RFID : Radio Fréquence Identification

M2M: MACHINE TO MACHINE

OWL: Web Ontology Language

RDF : Resource Description Framework

IAP : 'Institut d'Astrophysique de Paris

RATP : Régie autonome des transports parisiens

ISBN : L'International Standard Book Number

IEEE: Institute of Electrical and Electronics Engineers

ISO: Organisation internationale de normalisation

NFC : Near Field Communication

WAN: Wide area network

LAN: Line area network

RNIS: réseau numérique à intégration de services

TCP : Transmission Control Protoco

UDP : User Datagram Protocol

Glossaire

Http : Hypertexte Transfer Protocol

FTP: File Transfer Protocol

COAP: Constrained Application Protocol

6LoWPAN: IPv6 LoW Power wireless Area Networks

MQTT: MQ Telemetry Transport

XMPP: Extensible Messaging and Presence Protocol

BLE: Bluetooth Low Energy

OWASP: Open Web Application Security Project

Table des matières

Chapitre 1 : Big data

Chapitre 2: Cloud Computing

Chapitre 3 : internet des objets (IoT)

Chapitre 4 : Sécurité d'internet des objets IOT

Chapitre 5: travaux de recherche et proposition

Table des matières

Liste Des Figures

Liste Des tableaux

Introduction générale

L'Internet des objets est un concept concrétisant la vision de l'informatique ubiquitaire et elle qu'imaginée en 1991 par Mark Weiser [web 42], où la technologie s'efface peu à peu dans l'environnement des utilisateurs, intégrée naturellement à l'intérieur des objets du quotidien 1. La technologie n'est plus alors représentée par un objet unique, l'ordinateur personnel, mais se présente au contraire sous la forme d'appareils spécialisés et simples d'emploi, capables de communiquer au travers de plusieurs types de réseaux sans fil : liseuses numériques, télévisions et montres connectées, ordinateurs de bord, téléphones intelligents, etc.

À l'origine, le terme Internet des objets a été utilisé pour la première fois en 1999 par Kevin Ashton [web 43] pour décrire des objets équipés de puces d'identification par radiofréquence(ou puce RFID). Chaque objet, identifié de manière unique et universelle, peut alors être rattaché à un ensemble d'informations le concernant, ces dernières étant lisibles par d'autres machines. Caractéristiques, état courant et position sont alors autant de métadonnées échangées entre les objets, formant un nouveau réseau qui leur est dédié : l'Internet des objets.

L'Internet des objets, ou Internet of Things (IoT) en anglais, transformera l'ensemble de la société, y compris nous-mêmes. À première vue, cette affirmation peut paraître exagérée, mais pensez à l'impact qu'a déjà eu Internet sur l'enseignement, les communications, les entreprises, la science, les organismes publics et les hommes. Internet est sans nul doute l'une des inventions les plus importantes et les plus significatives de toute l'histoire de l'humanité.

Le concept a toutefois évolué avec le temps et s'est généralisé vers une approche consistant à connecter un très grand nombre d'objets du quotidien au réseau Internet, les dotant ainsi d'une identité propre et leur permettant, entre autres, d'offrir des services et de collecter des informations de manière autonome. L'objectif ambitieux derrière cette interconnexion est double, consistant en premier lieu dans la mise en place d'une infrastructure de communication machine à machine (M2M) à grande échelle de façon à permettre à ces machines de mieux « percevoir » le monde qui les entoure [Web 44]. En second lieu réside une volonté d'offrir les abstractions nécessaires aux êtres humains pour interagir avec ces machines, et par extension avec le monde physique, aussi simplement qu'avec le monde virtuel que nous connaissons aujourd'hui [Web 45]. En d'autres termes, les utilisateurs de l'Internet des objets devraient être en mesure de manipuler l'environnement physique de la même façon qu'ils manipulent aujourd'hui des abstractions de haut niveau telles que les fichiers et les dossiers, les pages Web et les hyperliens, ou encore les profils et les relations (p. ex. sur les réseaux sociaux).

Dites-vous maintenant que l'IoT représente la prochaine évolution d'Internet et permettra d'améliorer considérablement sa capacité à rassembler, à analyser et à restituer des

données que nous pourrons ensuite transformer en informations, en connaissances et enfin en savoir. Dans ce contexte, l'importance de l'IoT paraît évidente.

Des projets IoT déjà en cours promettent de combler les écarts de richesse, d'améliorer la distribution des ressources mondiales aux populations défavorisées et de nous aider à comprendre notre planète, ce qui nous permettra d'adopter un comportement plus proactif au lieu de simplement réagir aux événements. Toutefois, plusieurs obstacles menacent de ralentir le développement de l'IoT, notamment la transition vers le protocole IPv6, la mise en place de normes communes et le développement de sources d'énergie pour des millions, voire des milliards de minuscules capteurs.

Heureusement, grâce aux efforts conjugués des entreprises, des administrations publiques, des organismes de normalisation et des universités, l'IoT poursuivra sa progression. Ce document a pour objectif de vous informer de façon simple et claire sur ce sujet, et de vous expliquer que cette évolution a véritablement le potentiel de transformer tout ce que nous savons actuellement.

C'est dans ce contexte que ce mémoire s'inscrit puisqu'elle s'organise autour de la proposition de solutions qui découlent du constat des difficultés à réellement réaliser une intégration des technologies propice à l'éclosion d'architectures efficaces et novatrices pour l'Internet des objets. L'émergence de normalisation et de standardisation à tous niveaux offrira une stabilité propre à rassurer les différents acteurs du domaine, ouvrant la voie à des développements plus sûrs et pérennes, facilitant les collaborations, et proposant des solutions pour chacune des strates évoquées. L'extension de l'Internet à ces nouveaux réseaux d'objets contraints soulève la question de la pertinence et/ou de l'adaptation des protocoles applicatifs à l'œuvre, des approches de programmation et des architectures logicielles, capables d'à la fois offrir la continuité promise dans l'appellation "Internet" tout en respectant les spécificités de chacune des parties prenantes. Les applications proposées devront simultanément se fondre dans la palette des outils usuels que l'internaute manipule, et intégrer ces objets intelligents dans leurs interactions.

1. Problématiques posées par l'Internet des objets

Pour permettre à l'Internet des objets d'atteindre son régime maximal et concrétiser les scénarios qui en découlent, plusieurs aspects doivent être étudiés. En effet, les concepts prometteurs imaginés par la communauté scientifique nécessitent de résoudre un certain nombre de problématiques à savoir :

- Grande échelle
- Hétérogénéité des objets connectés
- Impact du monde physique (grands flux continus de données
- Variabilité de l'environnement
- Sécurité des biens et des personnes
- Respect de la vie privée des utilisateurs.

1.1. Grande échelle

De toutes les difficultés posées par l'Internet des objets, son échelle mondiale et le très grand nombre d'objets impliqués sont les plus importantes il paraît raisonnable, en ce qui concerne le réseau de communication global proprement dit, de considérer le réseau Internet comme un candidat prometteur. En effet, celui-ci est déjà massivement déployé à l'échelle mondiale et interconnecte plusieurs milliards de machines diverses et variées. En cela, Internet est le plus grand réseau connu jamais déployé, et paraît donc une bonne base pour l'Internet des objets. Par ailleurs, on estime qu'aujourd'hui une portion non négligeable des connexions sur Internet se font déjà entre objets connectés entre un entraînant ainsi un ensemble de problèmes qu'on peut résumer comme suit

1.1.1. Adressage et nommage

- Le très grand nombre d'objets nécessite
- Un espace d'adressage important
- Et augmente significativement la quantité d'informations que les serveurs de noms doivent stocker pour assurer leur rôle d'association entre les noms d'objet et leurs adresses.

1.1.2. Découverte

- Enregistrer et rechercher des objets par leur nom ou par leurs caractéristiques fondamentales pour la réalisation des scénarios où les objets ne sont pas connu à l'avance. Cependant, stocker et parcourir l'ensemble des objets de manière centralisé en 'est pas envisageable à une telle échelle.

1.1.3. Accès

- La multitude d'objets différents complexifie le rapatriement des données (data collection), tandis que les grands volumes de données complexifient leur traitement (datamining et data agrégation).

1.2. Hétérogénéité de l'Internet des objets

Conformément aux usages pour lesquels ils ont été conçus, les propriétés et les capacités des objets varient significativement, contribuant à faire de l'Internet des objets un écosystème certes riche, mais aussi hétérogène, les effets de cette hétérogénéité peuvent être résumés comme suit :

1.2.1. Hétérogénéité fonctionnelle

Les objets possèdent des capacités spécifiques (statique ou mobile, alimenté en continu ou par une batterie, ressources matérielles, capteurs et actionneurs disponibles, etc.) et à chacune d'elle correspond des contraintes particulières (connectivité intermittente, durée de vie, tâches réalisables, etc.). Différentes approches et techniques doivent être considérées pour gérer les objets en fonction de leurs contraintes et de leurs différences propres.

1.2.2. Hétérogénéité technique

Les technologies matérielles et logicielles utilisées pour construire les objets sont multiples et compromettent l'idéal de collaboration autonome entre objets. De plus, le développement d'applications est complexifié, nécessitant des développeurs des connaissances spécifiques sur le fonctionnement de chaque objet.

1 .3. Internet des objets et le monde physique

L'Internet des objets est fondamentalement influencé par les caractéristiques du monde physique et des outils qui permettent de le mesurer, comme résumé dans ce qui suit :

1.3.1. Flux de données

Les informations produites par les capteurs sont naturellement liées au temps, sous la forme de flux de mesures ou d'évènements. Cette particularité s'oppose radicalement aux approches classiques basées sur des ensembles de données finis, et nécessite une réflexion différente en ce qui concerne la représentation des données (data model) et leur traitement (computation model).

1.3.2. Capteurs

Les capteurs sont connus pour produire régulièrement des mesures erronées, imprécises ou incomplètes, sans qu'il soit possible de prédire précisément à quel moment celles-ci vont apparaître (transitent fautes) **[web 41]**. Des techniques spécifiques de détection, de correction ou d'atténuation d'erreurs doivent être mises en œuvre, tout particulièrement dans un environnement multi-capteur comme l'Internet des objets.

1.3.3. Variabilité

Le monde physique est un environnement changeant, spécifiquement dans le cadre des objets mobiles qui doivent composer avec leurs limites en énergie, la connectivité intermittente et les mouvements de leurs utilisateurs. Aussi, les objets doivent pouvoir s'adapter aux changements qui surviennent au cours du temps, conformément aux besoins des utilisateurs.

Chapitre 1

1. Introduction

Ces dernières années, un énorme buzz autour du big data s'est développé à tel point qu'un certain nombre d'analystes n'y ont vu qu'un phénomène marketing de plus, visant à favoriser les ventes des fournisseurs de technologie. Les protagonistes des systèmes d'information, à commencer par les entreprises utilisatrices, s'aperçoivent à présent que le phénomène est bien réel. Et les enjeux sont de taille, car le big data n'est pas qu'une question technique de volumétrie et de stockage. Il constitue au contraire l'opportunité de comprendre le contenu de ces nouvelles sources et d'entirer profit. La valeur des données non structurées issues des réseaux sociaux en est le parfait exemple.

Le sujet des BIG DATA est émergent partout dans le monde. La tendance est maintenant à l'utilisation des données hétérogènes pour acquérir de nouvelles connaissances, réaliser des gains de productivité, faciliter la prise de décision. Cela va de pair avec l'usage de technologies qui vont faciliter l'intégration, l'organisation, le traitement et la réutilisation de toutes ces données.

Le big data ne se réduit pas à une problématique unique ; il se décline au contraire en différentes variantes selon le métier et la situation particulière de chaque entreprise. Les techniques doivent donc être choisies et intégrées en fonction de la situation. Les projets big data ont cependant des caractéristiques communes dans leur approche et leur méthode de mise en œuvre.

Les BIG DATA vont bouleverser les habitudes. Sur Wikipédia, on en trouve une definition très intéressante qui precise un certain enjeu technologique : "***Big data consists of datasets that grow so large that they become awkward to work with using on-hand database management tools. Difficulties include capture, storage, search, sharing, analytics, and visualizing*****" [web1].**

En effet, il y a de grands volumes de données à exploiter, ceux-ci sont exponentiels. Ces données auront besoin d'être stockées pour ensuite être analysées et réutilisées mais pas avec les technologies traditionnelles.

2. Le phénomène BIG DATA

L'évolution du SI amène les entreprises à traiter de plus en plus de données issues de sources toujours plus variées.

Les prévisions de taux de croissance des volumes de données traitées dépassent les limites des technologies traditionnelles. On parle de péta octet (billiard d'octets) voir de zetta octet (trilliard d'octets). L'explosion quantitative des données numériques a obligé les chercheurs à trouver de nouvelles manières de voir et d'analyser le monde. Il s'agit de découvrir de nouveaux ordres de grandeur concernant la capture, la recherche, le partage, le stockage, l'analyse et la présentation des données. Ainsi est né le « Big Data ». **[1]**

2.1. Définition :

Littéralement, ces termes signifient méga données, grosses données ou encore données massives. Ils désignent un ensemble très volumineux de données qu'aucun outil classique de gestion de base de données ou de gestion de l'information ne peut vraiment travailler. En effet, nous procréons environ 2,5 trillions d'octets de données tous les jours. Ce sont les informations provenant de partout : messages que nous nous envoyons, vidéos que nous publions, informations climatiques, signaux GPS, enregistrements transactionnels d'achats en ligne et bien d'autres encore. Ces données sont baptisées Big Data ou volumes massifs de données. Les géants du Web, au premier rang desquels Yahoo (mais aussi Facebook et Google), ont été les tous premiers à déployer ce type de technologie.**[1]**

Le Big Data se caractérise par la problématique des 3V :

- **Vélocité :** la vitesse à laquelle les données sont traitées simultanément.
- **Variété** : l'origine variée des sources de données qui arrivent non-structurées (formats, codes, langages différents...).
- **Volume :** le poids total des données collectées

2.2. L'origine des données

Pour piloter son activité, l'entreprise doit enrichir ses données avec celles du Big Data (volumes de données sans limite, réponses en temps réel, personnalisation accrue,…). Le Big Data permet ainsi de faire passer l'entreprise de l'analyse reporting à l'analyse prescriptive. Tour d'horizon des quatre sources d'information sur lesquelles s'appuie le Big Data.

L'information produite par l'entreprise (journal des ventes, états des stocks, liste des clients et prospects…) s'organise dans des bases de données (dites de production), elles-mêmes agrégées dans des entrepôts de données (datawarehouse ou datamarts).

Ces données sont ensuite traitées sous forme de cubes décisionnels pour permettre de visualiser des indicateurs de performance sous différentes dimensions (temporelle, géographique, catégories de produits, segmentation client,…). **[1]**

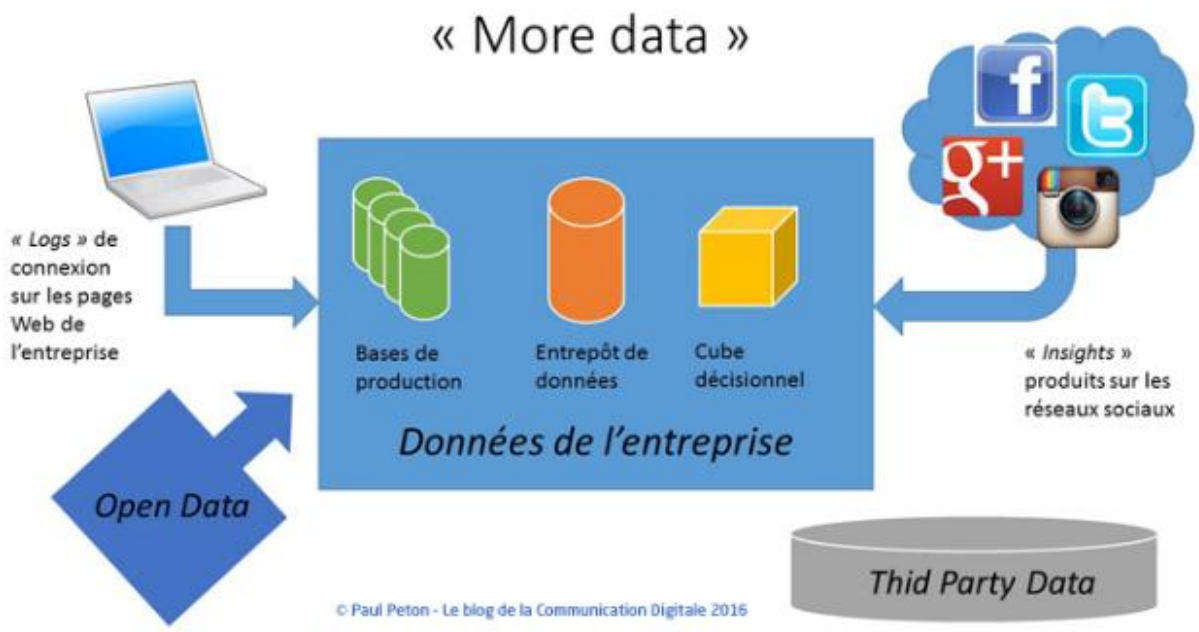

Figure 1.1 : Source de données Big Data [1]

Une approche Big Data permet d'enrichir les données de l'entreprise avec celles de sources externes. Il n'est pas ici question de tendre vers « l'infobésité » (en accumulant toujours plus d'information non exploitée) mais plutôt de se donner de nouveaux angles de vue sur l'activité de l'entreprise, la conjoncture dans son secteur, ou encore son positionnement sur le Web. [Web 2]

Le Big Data s'appuie sur quatre sources de données :

- Les « logs » des sites web
- Les « insights » des médias sociaux
- Les « third party data »
- L'Open data

2.2.1. Les « logs » (journaux de connexion) issus du trafic sur le site officiel de l'entreprise

Votre entreprise dispose certainement d'une vitrine sur le Web au travers de son site officiel. Ce site génère du trafic qu'il est indispensable d'analyser. Pour une approche plus fine, et donc plus riche en informations, on disposera des trackers sur les différentes pages afin de mesurer les chemins de navigation, ou encore les temps passés sur chaque page… Voire les déplacements de la souris sur l'écran !

D'autres questions intéressantes, et donc d'autres sources de données, sont les chemins pris par les visiteurs pour parvenir sur le site : moteurs de recherche, annuaires, rebonds depuis d'autres sites… **[1]**

Citons parmi les solutions d'analyse les plus connues : Google Analytics, Adobe Omniture, Coremetics.

2.2.2. Le contenu et les mesures de réputation (« insights ») issus des médias sociaux

Se définir une identité numérique, animer une communauté sont des pratiques maintenant bien ancrées. C'est une source de données, venant concurrencer les traditionnelles enquêtes par questionnaires.

Attention toutefois au travers des « mesures de vanité », très faciles à obtenir (like, partage, retweet…). Les signaux négatifs sont moins nombreux, mais expriment un geste fort de la part de leur auteur. Pensez donc à mesurer les publications masquées ou les désabonnements (et à y réagir !).

Une approche complémentaire, mêlant méthodes quantitatives et qualitatives, consiste à recueillir les commentaires aux publications et à y appliquer des algorithmes d'analyse de sentiment.

Quelques pistes pour suivre vos différents comptes : Hootsuite, Radian6 ou encore les API mises à disposition et interrogées avec le complément Power Query pour Excel, IRaMuTeQ pour l'analyse de données textuelles**. [1]**

2.2.3. La « third party data » : des données comportementales pour mieux cibler

Des acteurs spécialisés du Web vous aident à collecter de l'information sur vos clients ou prospects et à améliorer ainsi les campagnes de communication. Les données sur les internautes (third party data) sont récoltées par ces entreprises via des formulaires ou des cookies. Au-delà des classiques informations d'identité (sexe, âge, CSP…), il est maintenant beaucoup plus efficace de mesurer les comportements (navigation, configuration matérielle, temps passé sur les pages…).

Quelques acteurs du domaine de la third party data : Bluekai, Exelate, Weborama, Datalogix…

2.2.4. L'open data : les données ouvertes et réutilisables

Les données ouvertes et réutilisables ne sont pas encore légions même si une mission gouvernementale est très active sur le sujet. Manque de complétude, niveau de détail insuffisant, relative ancienneté sont les défauts actuels de nombreux jeux de données. Toutefois, c'est un champ d'investigation qu'il ne faut pas négliger, ne serait-ce que par son faible coût (celui du temps passé à chercher !) et son développement inéluctable.

2.3. Les principaux acteurs

Les fournisseurs historiques de solutions IT tels que HP, Oracle, IBM ou SAP figurent parmi les principaux acteurs du Big Data. Ainsi, IBM propose depuis fin 2011 InfoSphereBigInsights Basic pour IBM SmartCloud Enterprise. Cette version pouvant gérer 10To de données est accessible gratuitement aux utilisateurs de Linux. Cependant, BigInsights Enterprise est payant. De

son côté, Microsoft a privilégié l'utilisation du framework Hadoop en 2011 au détriment de LINQ to HPC. Le géant de l'informatique l'a ainsi utilisé pour développer Windows Azure et Windows Server. L'éditeur de Redmond a également développé SQL Server2012, un logiciel spécialisé dans la gestion des bases de données dans le souci de répondre aux besoins du BigData[web3].

2.4. Enjeux technologiques :

2.4.1. Stockage des données

Le premier élément structurant dans le contexte Big Data est le socle de stockage des données. Anciennement, la solution était les DatawareHouse (entrepôts de données), qui ont évolué pour supporter de plus grandes quantités de données et faire porter par le stockage, une capacité de traitement étendue. Les solutions de DatawareHouse ont toutes en commun un modèle de données profondément structuré (schéma, base de données, tables, types, vues, etc) et un langage de requête SQL.

Le Big Data vient rompre cette approche ; l'approche du Big Data consiste en 2 grands principes.

Premièrement, le principe de la scalabilité (horizontale) des clusters de traitement. Puis deuxièmement, on peut s'affranchir de certaines contraintes inhérentes aux bases de données relationnelles traditionnelles et qui ne sont pas forcément nécessaires pour le Big Data. C'est le cas de l'ACIDité (Atomicité, Cohérence, Isolation et Durabilité) **[2]**.

Pour mettre en œuvre cette approche avec une infrastructure simple, scalable (mot utilisé pour indiquer à quel point un système matériel ou logiciel parvient à répondre à une demande grandissante de la part des utilisateurs, il traduit aussi la capacité de montée en charge), du matériel à bas coût, le Framework Hadoop est utilisé pour la gestion du cluster, l'organisation et la manière de développer. La solution la plus emblématique de cette approche est Hadoop et son écosystème [3].

2.4.2. Big Data et Hadoop

Hadoop est un projet Open Source géré par Apache Software Fundation basé sur le principe MapReduce et de Google File System, deux produits Google Corp. Le produit est écrit en langage Java.

Hadoop peut être considéré comme un système de traitement de données évolutif pour le stockage et le traitement par lot de très grande quantité de données. Il est tout à fait adapté aux stockages de grande taille et aux analyses de type "ad hoc" sur de très grandes quantités de données. Le besoin en analyse de grandes masses de données devient toujours plus pressant. Les analyses des données collectées sur le web, les traces laissées par les clients et prospects sont les applications les plus souvent citées.

Hadoop permet d'exécuter des applications sur des systèmes en cluster dotés de milliers de nœuds impliquant des centaines de téraoctets de données. Son système de fichiers distribué favorise un taux élevé de transfert de données entre les nœuds et permet un fonctionnement ininterrompu du système en cas de défaillance d'un d'entre eux. Cette approche diminue le risque de panne majeure, même lorsqu'un nombre important de nœuds deviennent inopérants.

L'écosystème Apache Hadoop se compose du noyau Hadoop, de MapReduce, du système de fichiers distribué (HDFS) Hadoop et d'un certain nombre de projets associés, notamment Apache Hive, HBase et Zookeeper.

L'infrastructure Hadoop est utilisée par des acteurs majeurs - dont Google, Yahoo, etc. - pour des applications faisant intervenir des moteurs de recherche, de la publicité en ligne ou tout autre gestion des Big Data. **[4]**.

a.Hadoop Distributed File System (HDFS)

HDFS est un système de fichiers distribué qui donne un accès haute-performance aux données réparties dans des clusters Hadoop. Comme d'autres technologies liées à Hadoop, HDFS est devenu un outil clé pour gérer des pools de Big Data et supporter les applications analytiques. Un dispositif de stockage et d'accès à des fichiers

HDFS est généralement déployé sur des serveurs à bas coût dits de commodité. Les pannes sont ainsi fréquentes. Le système de fichiers est donc conçu pour être extrêmement tolérant aux pannes, tout en facilitant le transfert rapide de données entre les nœuds du système. Cela permet de maintenir les systèmes Hadoop opérationnels si un nœud tombe en panne. Ce mécanisme réduit le risque d'une défaillance critique, même si la panne concerne plusieurs nœuds.

Quand HDFS recueille une donnée, le système segmente l'information en plusieurs briques et les distribue sur plusieurs nœuds du cluster, ce qui permet alors le traitement en parallèle. Le système de fichiers copie chaque brique de donnée plusieurs fois et distribue les copies sur chacun des nœuds, plaçant au moins une copie sur un serveur séparé dans le cluster. Résultat, les données, stockées sur des nœuds défaillants, peuvent être retrouvées ailleurs dans le cluster. Le traitement peut se poursuivre en dépit de la panne.

HDFS est développé pour supporter les applications avec de grands volumes de données, comme les fichiers individuels dont la quantité peut se compter en teraoctets. Il s'adosse à une architecture maître / esclave, chaque cluster comprenant un NameNode unique en charge des opérations du système de fichiers. NameNode supporte également les DataNodes qui gèrent le stockage de données sur chaque nœud. **[4]**

- **Organisation des machines pour HDFS**

Un cluster HDFS est constitué de machines jouant différents rôles exclusifs entre eux :**[2]**

- ✓ L'une des machines est le maître HDFS, appelé **le namenode**. Cette machine contient tous les noms et blocs des fichiers, comme un gros annuaire téléphonique.
- ✓ Une autre machine est **le secondary namenode**, une sorte de namenode de secours, qui enregistre des sauvegardes de l'annuaire à intervalles réguliers.
- ✓ Certaines machines sont des clients. Ce sont des points d'accès au cluster pour s'y connecter et travailler.
- ✓ Toutes les autres machines sont des **datanodes**. Elles stockent les blocs du contenu des fichiers.

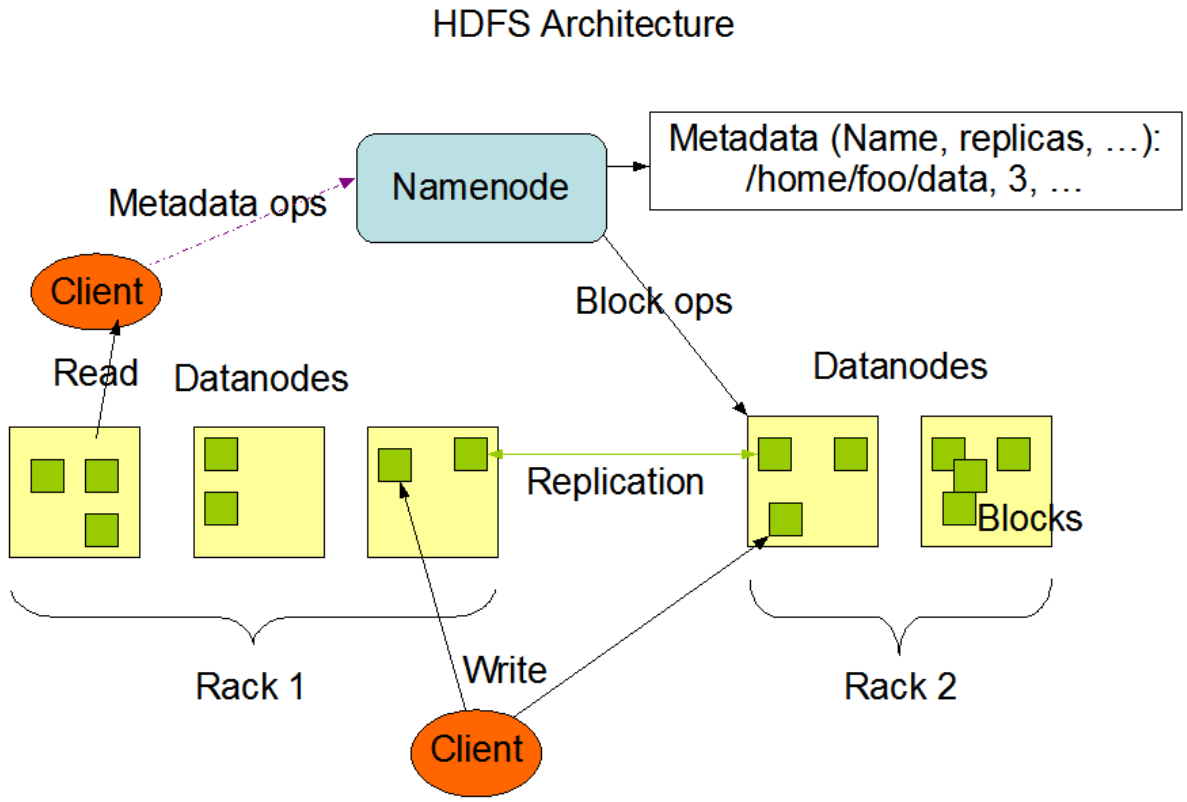

Figure 1.2 : Organisation des machines pour HDFS [2]

1. Namenode

Un Namenode est un service central (généralement appelé aussi maître) qui s'occupe de gérer l'état du système de fichiers. Il maintient l'arborescence du système de fichiers et les métadonnées de l'ensemble des fichiers et répertoires d'un système Haddon. Le Namenode a une connaissance des Datanodes (étudiés juste après) dans lesquels les blocs sont stockés. Ainsi, quand un client sollicite Hadoop pour récupérer un fichier, c'est via le Namenode que l'information est extraite. Ce Namenode va indiquer au client quels sont les Datanodes qui contiennent les blocs. Il ne reste plus au client qu'à récupérer les blocs souhaités.

Toutes ces métadonnées, hormis la position des blocs dans les Datanodes, sont stockées physiquement sur le disque système dans deux types de fichiers spécifiques edits_xxx et fsimage_xxx.

La connaissance de la position des blocs dans les Datanodes est reconstruite à chaque démarrage du Namenode dans un mode appelé safe mode. Pendant le safe mode, l'écriture sur HDFS est impossible, le Namenode charge les fichiers edits_xxx et fsimage_xxx et attend le retour des Datanodes sur la position des blocs. Une fois toutes les opérations réalisées, le safe mode est relâché et l'accès en écriture est de nouveau autorisé. Soyez patient sur la durée du safe mode. Celui-ci peut être très long si vous avez beaucoup de fichiers à traiter.

Comme vous l'aurez remarqué, le Namenode charge tout en mémoire. Cela devient donc problématique si vous avez énormément de petits fichiers à gérer.

Chaque fichier, répertoire et bloc dans HDFS est représenté comme un objet dans la mémoire et occupe 150 octets. Si, par exemple, vous avez 10 millions de fichiers à gérer, le Namenode devra disposer d'un minimum de 1,5 Go de mémoire. C'est donc un point important à prendre en compte lors du dimensionnement de votre cluster. Le Namenode est relativement gourmand en mémoire **[5]**.

2. Secondary Namenode

Le Namenode dans l'architecture Hadoop est un point unique de défaillance (Single Point of Failure en anglais). Si ce service est arrêté, il n'y a pas moyen de pouvoir extraire les blocs d'un fichier donné. Pour répondre à cette problématique, un Namenode secondaire appelé Secondary Namenode a été mis en place dans l'architecture Hadoop. Son fonctionnement est relativement simple puisque le Namenode secondaire vérifie périodiquement l'état du Namenode principal et copie les métadonnées via les fichiers edits_xxx et fsimage_xxx. Si le Namenode principal est indisponible, le Namenode secondaire prend sa place **[5]**.

3. Datanode

Précédemment, nous avons vu qu'un Datanode contient les blocs de données. Les Datanodes sont sous les ordres du Namenode et sont surnommés les Workers. Ils sont donc sollicités par les Namenodes lors des opérations de lecture et d'écriture. En lecture, les Datenodes vont transmettre au client les blocs correspondant au fichier à transmettre. En écriture, les Datanodes vont retourner l'emplacement des blocs fraîchement créés. Les Datanodes sont également sollicités lors de l'initialisation du Namenode et aussi de manière périodique, afin de retourner la liste des blocs stockés **[5]**.

b. Algorithme « MapReduce »

b.1 définition

MapReduce est un framework de développement informatique, introduit par Google, dans lequel sont effectués des calculs parallèles, et souvent distribués, de données potentiellement très

volumineuses (> 1 terabyte).Les terminologies de « Map » et « Reduce », et leurs idées générales, sont empruntées aux langages de programmation fonctionnelle utilisés pour leur construction (map et réduction de la programmation fonctionnelle et des langages de programmation tableau).**[6]**

MapReduce s'articule en deux étapes :

- Dans l'étape Map le nœud analyse un problème, le découpe en sous-problèmes, et les délègue à d'autre nœuds (qui peuvent en faire de même récursivement). Les sous-problèmes sont ensuite traités par les différents nœuds à l'aide de la fonction Map qui a un couple (clé, valeur) associé à un ensemble de nouveaux couples (clé, valeur).
- Vient ensuite l'étape Reduce, où les nœuds les plus bas font remonter leurs résultats au nœud parent qui les avait sollicités. Celui-ci calcule un résultat partiel à l'aide de la fonction Reduce (réduction) qui associe toutes les valeurs correspondantes à la même clé à une unique paire (clé, valeur). Puis il remonte l'information à son tour.

À la fin du processus, le nœud d'origine peut recomposer une réponse au problème qui lui avait été soumis. Il y a plusieurs implémentations de ce framework dans différents langages (C++, Java, Pytho,..etc.) et par de nombreux organismes **[web4]**.

b.2. Principe général de MapReduce

Les concepts de Map et de Reduce ne sont pas nouveaux puisqu'ils ont été empruntés aux langages fonctionnels, sauf que Google les a efficacement propulsés dans l'univers du calcul distribué et du grand volume de données. Ils sont utilisés pour implémenter des opérations de base sur les données comme le tri, le filtrage, la projection, l'agrégation ou le regroupement. **[6]**

Pour expliquer les concepts de Map et de Reduce, partons de l'exemple du compteur de mots fréquemment utilisés, avec une légère variante.

Voiture la-le elle de elle la se la maison voiture

Tous les mots sont comptabilisés à l'exception du mot « se ».

On distingue clairement sur la première ligne que le mot la est répété trois fois et que le mot voiture est répété deux fois.

La figure ci-dessous énumère les différentes étapes qui seront présentées par la suite.

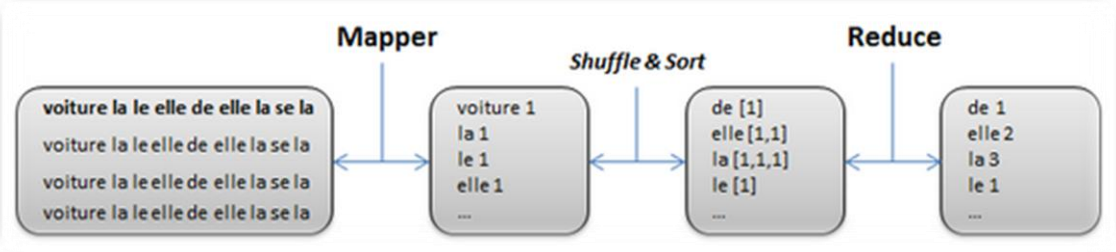

Figure 1.3 : Les différentes étapes de MapReduce

En Résumé :

- La fonction Map s'écrit de la manière suivante : map (clé1, valeur1) → List (clé2, valeur2). À partir d'un couple clé/valeur, la fonction map retourne un ensemble de nouveaux couples clé/valeur. Cet ensemble peut être vide, d'une cardinalité un ou plusieurs. Dans notre exemple, la clé d'entrée correspond au numéro de ligne dans le fichier et la valeur vaut voiture la-le elle de elle la se la maison voiture.
- La fonction reduce s'écrit de la manière suivante : reduce (clé2, List (valeur2)) → List(valeur2). À partir des groupes de valeurs associées à une clé, la fonction reduce retourne généralement une valeur ou aucune, bien qu'il soit possible de retourner plus d'une valeur. Suite à l'appel de la fonction reduce, le résultat de l'exemple est le suivant.

Le résultat des fonctions map et reduce est donné ci-dessous **[7]**.

Retour de la fonction map	Retour de la fonction reduce
(voiture, 1) (la, 1) (le, 1) (elle, 1) (de, 1) (elle, 1) (la, 1) (la, 1) (maison, 1) (voiture, 1)	**(de, 1)** **(elle, 2)** **(la, 3)** **(le, 1)** **(maison, 1)** **(voiture, 2)**

Tableau 1.1 : Retour de la fonction Map et de la fonction Reduce

c. HADOOP YARN

Yarn « YetAnother Resource Negotiator » est une technologie qui gère l'utilisation des ressources dans un « Cluster ». Depuis la version 2.0 de Hadoop, Yarn est plus générale que MapReduce et propose une nouvelle architecture de la fonction « Jobtracker » qui consiste à séparer ses deux tâches de gestion de ressources et celle de planification et surveillance des travaux. Cela permet d'exécuter des tâches qui ne se basent pas sur MapReduce comme Spark ainsi que des tâches MapReduce sur le même « Cluster ». **[web3]**

d. HADOOP COMMUN

Hadoop Common (anciennement HadoopCore) est un ensemble de composants communs pour la gestion des systèmes de fichiers distribués (sérialisation, Java RPC, etc.). Il est notamment utilisé par les sous-projets MapReduce et *HDFS*. **[web3]**

2.4.3. NO SQL

On ne peut pas parler de Big Data sans citer le NOSQL, Not Only SQL. Il est venu pour solutionner les difficultés rencontrées pendant la gestion des données classées « Big Data » avec les systèmes SGBD relationnelles. **[8]**

a. Problématique

Suite à des besoins de haute disponibilité, d'accès, de calculs des volumes impressionnants de données pour des grandes compagnies telles que Google, Amazon, Facebook, Twitter, LinkedIn , les solutions SQL offertes dans le marché, ne répondaient pas aux besoins, plus car leur limites ont été atteintes.**[9]**

b. Les limites :

- Volumétrie.
- Difficulté de la mise à jour sans cesse de serveur de plus en plus puissant.
- Analyser des quantités de données faramineuses (Peta octets).
- Tolérance aux pannes (serveur maître tombe)… etc.

D'où l'apparition de solutions basées sur des systèmes distribués qui sont les SGBD NOSQL. **[9]**

c. Historique du mouvement NOSQL

En 1998, le monde entend pour la première fois le terme NOSQL. Terme inventé et employé par Carlo Strozzi pour nommer son SGBD relationnel Open Source léger qui n'utilisait pas le langage SQL. Ironiquement, le trouvail de M. Strozzi n'a rien à voir avec la mouvance NOSQL que l'on connaît aujourd'hui, vu que son SGBD est de type relationnel. En effet, c'est en 2009, lors d'un

rassemblement de la communauté des développeurs des SGBD non-relationnels, que le terme NOSQL a été mis au goût du jour pour englober tous les SGBD de type non-relationnel. **[9]**

d. Définition

Le NOSQL est un type de base de données, c'est une manière de stocker et de récupérer des données de façon rapide, un peu comme une base de données relationnelle, sauf qu'il n'est pas basé sur des relations mathématiques entre les tables comme dans une base de données relationnelle traditionnelle. **[10]**

e. Comment le NOSQL fonctionne ?

Scalabilité est l'aptitude d'un système à conserver, maintenir son niveau de performance par augmentation des ressources matérielles On distingue deux types de scalabilité :

- Verticale ou interne : ajout de RAM, processeur au sein d'une machine ou remplacement par une de plus grand gabarit.
- Horizontale ou externe : Le principe consiste à simplement rajouter des serveurs identiques en parallèle afin de répondre à l'augmentation de la charge. **[10]**

Le but des systèmes NOSQL est de renforcer la scalabilité horizontale.

f. Type de base de données NOSQL

Il en existe 4 types distincts qui s'utilisent différemment et qui se prêtent mieux selon le type de données que l'on souhaite y stocker :

f.1. Les bases de données clé-valeur

Il s'agit de la catégorie de base de données la plus basique. Dans ce modèle chaque objet est identifié par une clé unique qui constitue la seule manière d'y accéder. Dans ce modèle on ne dispose généralement que des quatre opérations de base : Create, Retrieve, Update, Delete.

Ces bases ont l'avantage d'être très performantes en lecture et en écriture et permettent une extensibilité (*scalability)* élevée. On les retrouve très souvent comme système de stockage de cache ou de sessions distribuées, notamment là où l'intégrité relationnelle des données est non significative. **[10]**

Exemple de base « clé/valeur » : SimpleDB, DynamoDB (Amazon), Voldemort (LinkedIn), Redis

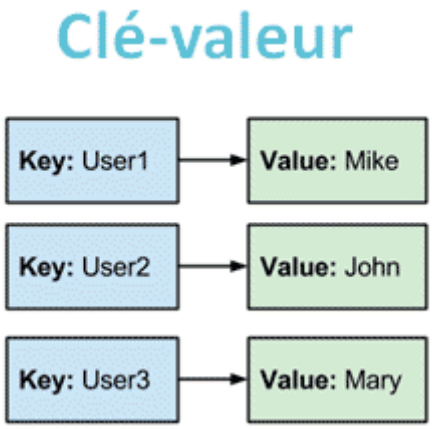

Figure 1.4 : Exemple base de données clé-valeur [10]

f.2 Les bases de données orientées document

La représentation en document est particulièrement adaptée au monde du Web. Il s'agit d'une extension du concept de clé-valeur qui représente la valeur sous la forme d'un document contient des données organisées de manière hiérarchique à l'image de ce que permettent XML ou JSON. Étant consciente du contenu qu'elle stocke, la base de données peut alors effectuer des indexations de différents champs et offrir des requêtes plus élaborées. **[web3]**

Les principaux avantages de ce type de système sont donc :

- Il est plus performant d'extraire des données pour une densité importante d'informations.
- Améliore grandement les performances sur les tris ou agrégations de données car ces opérations sont réalisées via des clés de lignes déjà triées. **[10]**

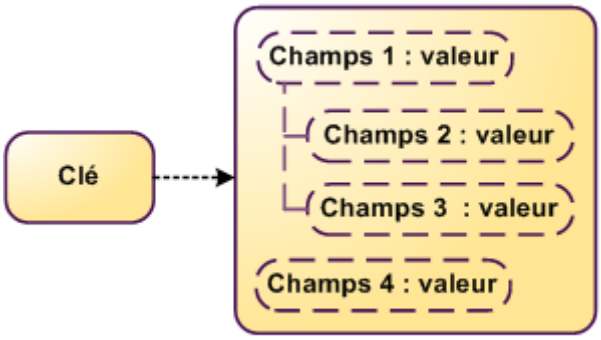

Figure 1.5 : Exemple base de données orientées document [10]

f.3. Les bases de données orientées colonnes

Le concept de colonnes est le plus simple à saisir, car l'analogie avec les bases relationnelles est proche. Dans les concepts à appréhender il existe des tables, ce qui permet de bien comprendre comment les données sont organisées. **[10]**

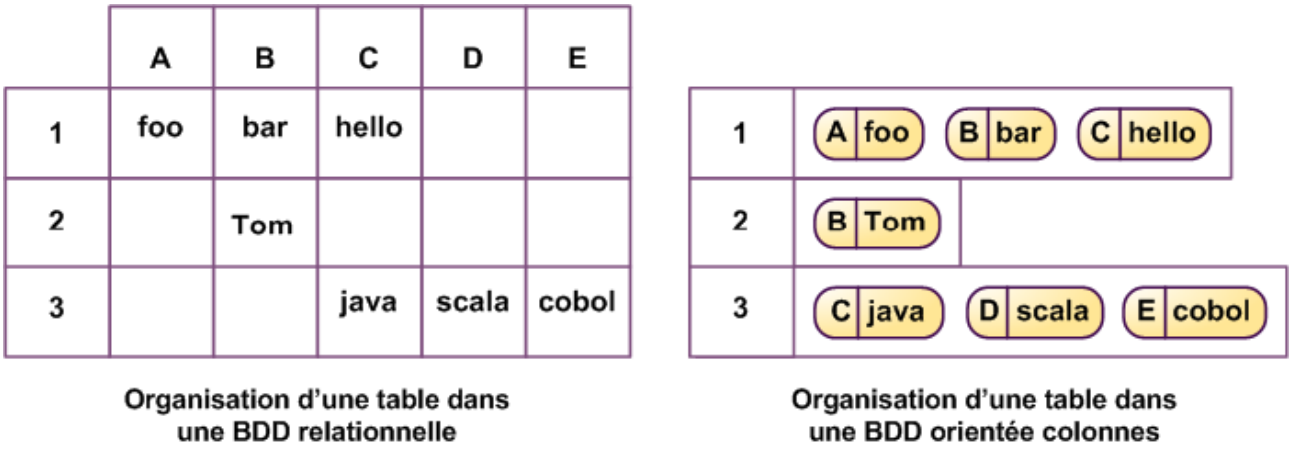

Figure 1.6 : Exemple base de données orientées colonnes [10]

f.4. Les bases de données orientées graphe

Les bases orientée graphe sont les moins connues de la mouvance NOSQL. Ces bases permettent la modélisation, le stockage ainsi que le traitement de données complexes reliées par des relations. **[10]**

Ce modèle est composé d'un :

- Moteur de stockage pour les objets : c'est une base ou chaque entité est nommée nœud.
- Mécanisme qui décrit les arcs : c'est les relations entre les objets, elles contiennent des propriétés de type simple (integer, string, date, ...).

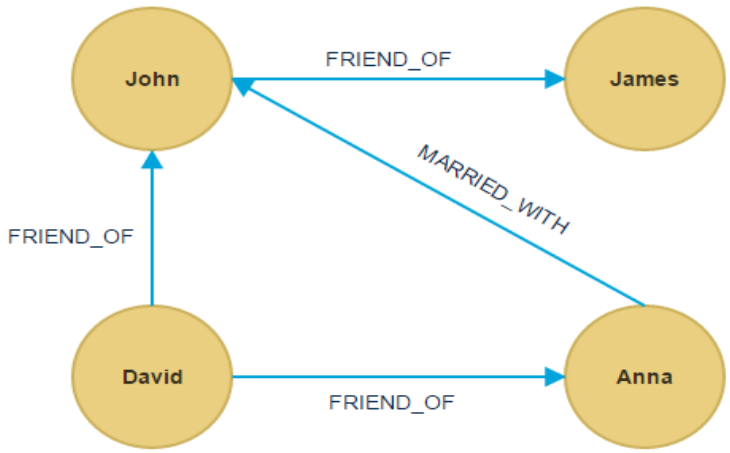

Figure 1.7 : Exemple base de données orientées graphe [10]

Conclusion

Nous avons abordé dans ce premier chapitre les principes des Big Data, ces caractéristiques, son fonctionnement ainsi que les différents domaines dans lesquels elles sont utilisées.

On a aussi recensé les différents modèles de bases de données NOSQL (Not Only SQL) qui existent, actuellement, dans le marché en accentuant sur les solutions les plus populaires.

Le Big Data ne constitue sans doute pas une révolution de l'ampleur de celles de l'agriculture ou de l'industrie, et à même d'engendrer un nouvel âge d'or. Cependant, en ouvrant à l'induction des données, il permet des transformations de nombreux champs d'activité : politique, social, éducatif, judiciaire, sportif, personnel… avec des aspects juridiques, éthiques ou encore psychologiques à ne pas minorer.

Par ailleurs, l'analyse de grands volumes de données défie également les moteurs de bases de données traditionnels. C'est pour répondre à ces différentes problématiques que sont nées les bases de données NOSQL, sous l'impulsion de grands acteurs du Web comme Facebook ou Google, qui les avaient développées à l'origine pour leurs besoins propres. Grâce à leur flexibilité et leur souplesse, ces bases non relationnelles permettent en effet de gérer de gros volumes de données hétérogènes sur un ensemble de serveurs de stockage distribués, avec une capacité de montée en charge très élevée.

Chapitre 2

1. Introduction

Le développement remarquable du Cloud Computing, ces dernières années, suscite de plus en plus l'intérêt des différents utilisateurs d'Internet et de l'informatique qui cherchent à profiter au mieux des services et des applications disponibles en ligne à travers le web en mode services à la demande et facturation à l'usage.

La disponibilité des services en ligne donne aussi la possibilité de ne plus s'approprier d'équipements informatiques mais de payer les frais en fonction de l'utilisation des ressources. Ce modèle attire déjà un grand nombre d'entreprises notamment les petites et moyennes entreprises « PME » et les très petites entreprises « TPE ».

Ce modèle informatique offre également la modularité des ressources informatiques (hard et soft) et leur disponibilité, en termes de volume et dans le temps, selon les besoins du client et à sa demande. Dans un contexte économique où les entreprises cherchent à rentabiliser au maximum les investissements et à limiter les coûts d'exploitation, le Cloud Computing se présente comme étant la solution de demain.

Le cloud computing est un concept assez récent dans le domaine de l'informatique. La première énonciation de ce concept date de 1960 par John McCarthy [Zhang, Cheng, Boutaba, 2010]. Il envisageait à cette époque que les matériels, équipements, ou installations informatiques pouvaient être délivrés aux utilisateurs sous forme de services. Ce concept a évolué dans le temps et est actuellement à une étape de concrétisation. **[11]**

2. Qu'est-ce que le cloud computing?

L'informatique dans le nuage est plus connue sous sa forme anglo-saxonne : « Cloud Computing », mais il existe de nombreux synonymes francophones tels que : « informatique dans les nuages », « infonuagique » (Québec) ou encore « informatique dématérialisée ». [Laurent, A. 2011]. **[web5]**

Même si les experts ne sont pas d'accords sur sa définition exacte, la plupart s'accordent à dire qu'elle inclue la notion de services disponibles à la demande, extensibles à volonté et à distance ou sur le net. En contradiction avec les systèmes actuels, les services sont virtuels et illimités et les détails des infrastructures physiques sur lesquels les applications reposent ne sont plus du ressort de l'utilisateur.

Selon le National Institute of Standards and Technology, le cloud computing englobe trois caractéristiques clés :

1. La mutualisation, de la part du fournisseur, de ressources éclatées ;
2. Des ressources accessibles en réseau ;
3. Des ressources accessibles rapidement, à la demande et de façon souple ;

Par exemple quelques définitions qui ont circulés :

« **Le cloud computing** est un modèle qui permet un accès réseau à la demande et pratique à un pool partagé des ressources informatiques configurables (telles que réseaux, serveurs, stockage, applications et services) qui peuvent être provisionnées rapidement et distribuées avec un minimum de gestion ou d'interaction avec le fournisseur de services. **»[Laurent, A. 2011].[web5]**

« **Le Cloud Computing** est une plateforme de mutualisation informatique fournissant aux entreprises des services à la demande avec l'illusion d'une infinité des ressources ». **[web6]**

Un des points essentiels de ces définitions est la notion de « scalability » ; d'extensibilité à la demande, d'élasticité, c'est à dire qu'on ne paie que ce qu'on utilise. C'est un avantage considérable par rapport à une infrastructure propre à l'entreprise où les serveurs sont très souvent sous-utilisés. On devrait avoir ici pas mal de références ?!

*« Donc le **Cloud Computing** est un concept qui consiste à déporter sur des serveurs distants des stockages et des traitements informatiques traditionnellement localisés sur des serveurs locaux ou sur le poste de l'utilisateur. Il consiste à proposer des services informatiques sous forme de service à la demande, accessible de n'importe où, n'importe quand et par n'importe qui* **». [12]**

L'idée principale à retenir est que le Cloud n'est pas un ensemble de technologies, mais un modèle de fourniture, de gestion et de consommation des services et des ressources informatiques localisés dans des Datacenter.

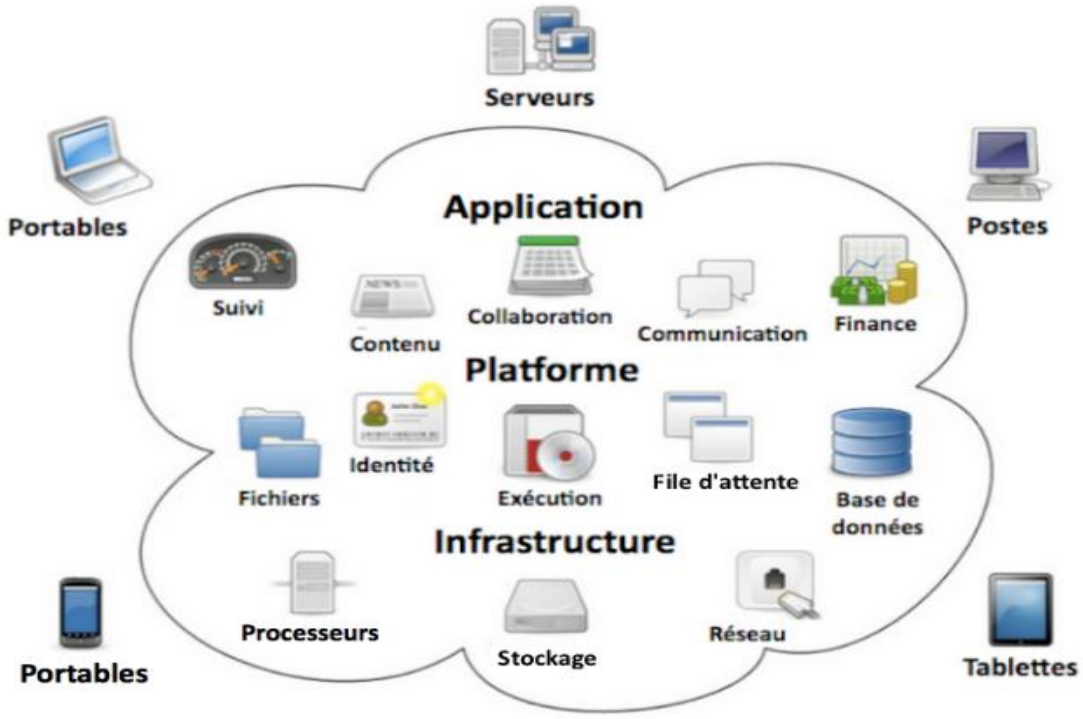

Figure 2.1 : Cloud Computing [12]

3. Services offerts par le cloud

Le cloud computing peut être décomposé en trois couches :

1. Application (SaaS, Software as a Service)
2. Platform (PaaS, Platform as a Service)
3. Infrastructure (IaaS, Infrastructure as a Service)

L'infrastructure as a Service (Iaas), est plutôt gérée par les architectes réseaux, la couche PaaS est destinée aux développeurs d'applications et finalement le logicielle comme un service (SaaS) est le « produit final » pour les utilisateurs. **[web7]**

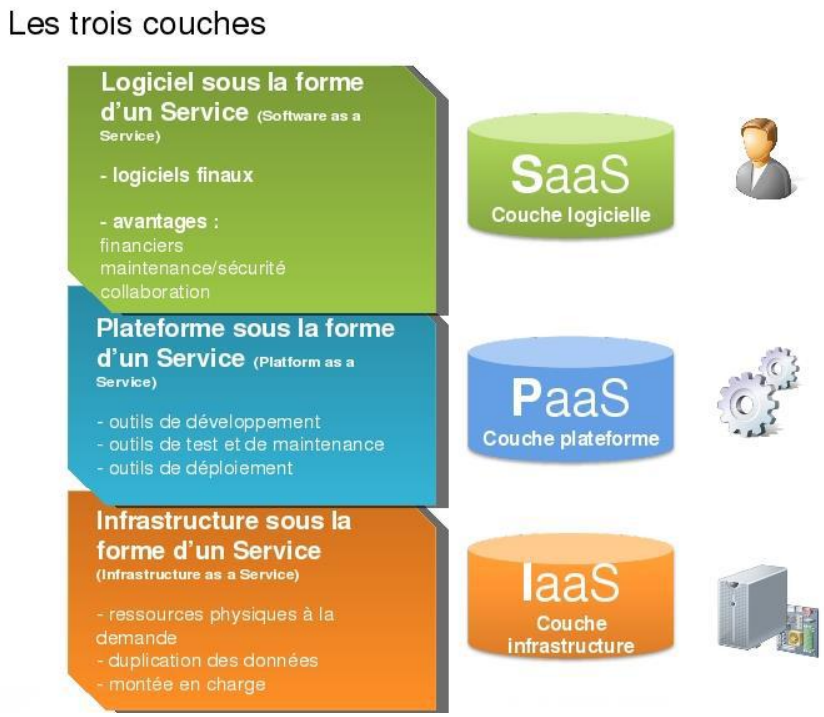

Figure 2.2 : Les couches du cloud computing [web7]

3.1. Le SaaS (Software as a Service)

Il s'agit d'un modèle de déploiement de logiciels par lequel un éditeur offre à ses clients la licence d'utilisation d'une application à la demande. Le logiciel est consommé comme un service. Les fournisseurs de logiciels en régime **SAAS** assurent soit l'hébergement de l'application sur leurs propres serveurs web, soit le téléchargement du logiciel sur l'environnement client (ils peuvent le désactiver après utilisation ou expiration du contrat). Les fonctionnalités à la demande peuvent être gérées en interne ou par un tiers fournisseur de services d'applications.

Les solutions SaaS constituent la forme la plus répandue de Cloud Computing. Les prestataires de solutions SaaS les plus connus sont Microsoft – offre Office365 (outils collaboratifs) Google – offre Google Apps (messagerie et bureautique).**[web7]**

3.2. Le PaaS (Plateforme as a Service)

Il consiste à fournir une plateforme informatique ainsi qu'une pile de solutions de développement, de test et d'exécution, le tout étant consommé comme un service. Cela facilite le déploiement d'applications, en réduisant le coût et la complexité généralement associés à l'achat et à la gestion des couches de base du matériel informatique et des logiciels. Toutes les fonctionnalités requises pour soutenir le cycle de vie complet des applications sont fournies.

Les principaux fournisseurs de PaaS sont : Microsoft avec AZURE, Google avec Google App Engine et Orange Business Services. **[web7]**

3.3. IaaS ou (Infrastructure as a Service)

Il consiste à fournir de l'infrastructure informatique (typiquement via une plateforme virtualité) comme un service. Plutôt que d'acheter des serveurs, des logiciels, de l'espace pour le centre de données ou de l'équipement réseau, les clients acquièrent ces ressources informatiques comme un service complètement externalisé. Ce service est facturé en fonction du niveau de consommation de la ressource. C'est une évolution de l'hébergement web.

L'IaaS offre une grande flexibilité, avec une administration à distance, et permet d'installer tout type de logiciel. En revanche, cette solution nécessite la présence d'un administrateur système au sein de l'entreprise, comme pour les solutions serveur classiques. Parmi les prestataires d'IaaS, on peut citer : Amazon avec EC2 ou Orange Business Services avec Flexible Computing.

En résume, Le concept d'informatique dans les nuages pourrait être également considéré comme de l'utility computing par lesquelles utilisateurs finaux peuvent accéder aux ressources informatiques presque aussi facilement,**[web7]**

4. Modèle de déploiement

Il peut être **privé** (intégré au système informatique de la société, à l'intérieur de son pare-feu) ou public. Dans sa forme externe (privé ou publique), la technologie « cloud » peut être considérée comme un nouveau moyen d'externaliser certaines parties du système informatique d'une société. Ce modèle est orienté vers les dépenses en investissements et vers une automatisation complète du centre de données**. [web7]**

4.1. Le nuage privé

L'infrastructure d'un nuage privé n'est utilisée que par un unique client. Elle peut être gérée par ce client ou par un prestataire de service et peut être située dans les locaux de l'entreprise cliente ou bien chez le prestataire, le cas échéant. L'utilisation d'un nuage privé permet de garantir, par exemple, que les ressources matérielles allouées ne seront jamais partagées par deux clients différents. **[13]**

4.2. Le nuage communautaire

L'infrastructure d'un nuage communautaire est partagée par plusieurs organisations indépendantes et est utilisée par une communauté qui est organisée au tour des mêmes besoins, vis-à-vis de son utilisation. Par exemple, dans le projet Open Cirrus, le nuage communautaire est partagé par plusieurs universités dans le cadre d'un projet scientifique commun. Son infrastructure peut être gérée par les organisations de la communauté qui l'utilise ou par un tiers et peut être située, soit au sein des dites organisations, soit chez un prestataire de service. **[13]**

4.3. Les nuages publics

L'infrastructure d'un nuage public est accessible publiquement ou pour un large groupe industriel. Son propriétaire est une entreprise qui vend de l'informatique en tant que service. **[13]**

4.4. Le nuage hybride

L'infrastructure d'un nuage hybride est une composition de deux ou trois des types de nuages précédemment cités. Les différents nuages qui la composent restent des entités indépendantes à part entière, mais sont reliés par des standards ou par des technologies propriétaires qui permettent la portabilité des applications déployées sur les différents nuages. Une utilisation type de nuage hybride est la répartition de charge entre plusieurs nuages pendant les pics du taux d'utilisation. **[13]**

5. La virtualisation

La virtualisation consiste à faire fonctionner un ou plusieurs systèmes d'exploitation sur un ou plusieurs ordinateurs. Cela peut sembler étrange d'installer deux systèmes d'exploitation sur une machine conçue pour en accueillir qu'un, mais comme nous le verrons par la suite, cette technique a de nombreux avantages. **[web5]**

Il est courant pour des entreprises de posséder de nombreux serveurs, tels que les serveurs de mail, de nom de domaine, de stockage pour ne citer que ceux-ci. Dans un contexte économique où il est important de rentabiliser tous les investissements, acheter plusieurs machines physiques pour héberger plusieurs serveurs n'est pas judicieux. De plus, une machine fonctionnant à 15 pour cent ne consomme pas plus d'énergie qu'une machine fonctionnant à 90 pour cent. Ainsi, regrouper ces serveurs sur une même machine peut donc s'avérer rentable si leurs pointes de charge ne coïncident pas systématiquement. **[web7]**

Enfin, la virtualisation des serveurs permet un bien plus grande modularité dans la répartition des charges et la reconfiguration des serveurs en cas d'évolution ou de défaillance momentanée.

Les intérêts de la virtualisation sont multiples. On peut citer :

1. L'utilisation optimale des ressources d'un parc de machines (répartition des machines virtuelles sur les machines physiques en fonction des charges respective)
2. L'économie sur le matériel (consommation électrique, entretien physique, surveillance)

6. Sécurité dans le Cloud Computing

La sécurité et la conformité émergent systématiquement comme les principales préoccupations des responsables informatiques lorsqu'il est question de Cloud Computing, des préoccupations encore plus accentuées lorsqu'il s'agit d'un Cloud public. La sécurité permet de garantir la confidentialité, l'intégrité, l'authenticité et la disponibilité des informations.**[13]**

Certaines questions légitimes reviennent sans cesse : **[14]**

1. Mes données sont-elles sures dans le Cloud ?
2. Où sont stockées mes données ?
3. Qui va avoir accès à mes données ?
4. Aurais-je accès à mes données à n'importe quel moment ?
5. Que deviendrons mes données s'il y a interruption du service ?

La mise sur pied d'une solution de Cloud Computing comporte des problèmes de sécurité inhérents à la solution elle -même. Le fait de centraliser toutes les informations sur un site pose un grand nombre de problèmes. On peut citer comme problème potentiel :

1. Une possible interruption massive du service.
2. Une cible de choix pour les hackers
3. Interface et API non sécurisé

Ce point de vulnérabilité du Cloud Computing fait l'objet depuis quelques années l'objet de recherches avancées. Il a été créé un organisme chargé de mettre sur pied des normes en matière de sécurité dans le Cloud Computing. Cet organisme s'appelle CSA (Cloud Security Alliance). Du travail de cet organisme, il en est ressorti certaines techniques utilisées de nos jours pour améliorer la sécurité du Cloud Computing.

7. L'avenir de l'informatique en nuage :

1. Il est difficile de déterminer si les types de services d'informatique en nuage vont profondément évoluer à court terme. Cependant, il est probable que leur disponibilité et leur capacité continuent d'augmenter, puisque les économies d'échelle favorisent le recours à des centres de données toujours plus grands, ce qui entraînera une migration vers des sites où le coût de l'énergie est moindre.
2. Alors que certains services pourraient bien migrer vers un nuage public en raison des économies potentielles, d'autres resteront dans un environnement privé, cardans de nombreux cas, le recours intelligent à des solutions à petite échelle est tout aussi efficace, voire plus efficace, que l'adoption de solutions à grande échelle.
3. Les aspects tels que la sécurité et le respect de la vie privée pourraient ralentir l'essor de l'informatique en nuage, car si les utilisateurs professionnels ou les autorités publiques ne font pas confiance aux nuages publics ou n'ont pas d'éléments prouvant qu'ils sont fiables, ils pourraient ne pas adopter ce modèle.
4. Cependant, le manque de concurrence, principalement dû à l'interopérabilité insuffisante, pourrait être l'un des obstacles majeurs à surmonter pour développer l’informatique en nuage.
5. À l'avenir, une difficulté importante consistera à définir les domaines possibles pour coordonner le développement de l'informatique en nuage au niveau européen afin d'éviter les doublons et le gaspillage. Cela supposerait de résoudre les problèmes majeurs en matière de normes pour garantir l'interopérabilité et de garantir une concurrence effective entre les fournisseurs, en s'attelant, par exemple, à l'intégration verticale des fournisseurs de services ou en facilitant la passation de marchés publics dédiés aux services d'informatique en nuage

afin d'encourager l'arrivée de nouveaux concurrents sur le marché et de coordonner les initiatives européennes et nationales dans ce domaine.

6. La connectivité va devenir un aspect de plus en plus important puisque l'essor de l'utilisation des services d'informatique en nuage va entraîner une plus grande dépendance des clients vis-à-vis des connexions à large bande à haut débit (y compris des réseaux mobiles sans fil de quatrième génération ou des autres technologies disponibles). **[11]**

Conclusion

De l'informatique utilitaire des années 60, au service bureau des années 70, tout en passant par l'émergence d'Internet et des avancées de virtualisation, le Cloud Computing comme les chiffres nous le confirme, est promis à un bel avenir.

L'objectif global de ce chapitre était de décrire les principaux enjeux du cloud computing en entreprise. Le but était d'analyser les différents points clés du Cloud et de vérifier que le concept n'est pas juste un phénomène de mode passager.

Le concept général du cloud computing devrait maintenant être clair. S'il fallait résumer le cloud computing en une phrase se serait un concept qui consiste à délocaliser au travers d'internet ses unités stockage et de calcul sur des serveurs gérés par une entreprise spécialisée afin de ne payer que les ressources que l'on a utilisées.

Chapitre 3

1. Introduction

L'Internet des objets est une concrétisation technique de l'informatique ubiquitaire où la technologie est intégrée naturellement aux objets du quotidien. Très prometteur, ce concept ouvre la voie vers une multitude de scénarios basés sur l'interconnexion entre le monde physique et le monde virtuel : domotique, e-santé, ville intelligente, logistique, sécurité, etc. Cependant, comme d'autres concepts prometteurs, celui-ci fait face à un certain nombre de problématiques techniques et non techniques qui nécessitent d'être étudiées pour permettre à l'Internet des objets d'atteindre son plein potentiel.

Jusqu'à présent, le nouveau paradigme qu'est l'Internet des objets émerge des travaux issus de plusieurs communautés scientifiques : les réseaux de capteurs et d'actionneurs sans fil, le Web, le cloud computing, l'identification par radio fréquence (radio-frequency identication, ou RFID) ou encore là la communication en champ proche (nearfield communication, ou NFC).

Le concept a toutefois évolué avec le temps et s'est généralisé vers une approche consistant à connecter un très grand nombre d'objets du quotidien au réseau Internet, les dotant ainsi d'une identité propre et leur permettant, entre autres, d'offrir des services et de collecter des informations de manière autonome. L'objectif ambitieux derrière cette interconnexion est double, consistant en premier lieu dans la mise en place d'une infrastructure de communication machine à machine (M2M) à grande échelle de façon à permettre à ces machines de mieux « percevoir » le monde qui les entoure.

2. Web 3.0

2.1 Historique du web

Le web est sans nul doute une technologie majeure du 21ème siècle. Et si sa nature, sa structure et son utilisation ont évolué au cours du temps, force est de constater que cette évolution a également profondément modifié nos pratiques commerciales et sociales.

Au début était le web, tout bête, ce que l'on nomme aujourd'hui le web 1.0 qui comprenait des pages statiques, on prenait de la communication papier et on la transférait sous forme numérique dans des pages html qui n'étaient pas souvent mises sinon jamais. Ce mouvement ne s'est toujours pas arrêté et l'on voit encore aujourd'hui des sites internet dits professionnels qui ne comportent qu'une plaquette scannée avec 2 ou trois lignes de texte, il s'agit de reproduire les modèles connus de l'édition papier, de les adapter aux navigateurs internet.

Le web 2.0 est apparu dans l'an 2003, Il est attribuable aux internautes. En effet, ces derniers désirent communiquer, partager et construire leurs réseaux. Internet se transforme en espace ouvert et collaboratif. Facebook, Twitter, YouTube deviennent des médias puissants, complémentaires aux médias dits traditionnels.

Le web 3.0 Il a commencé à apparaître en 2008, Cette phase est définie par l'intégration du web sémantique. Fondé sur les langages standards tels que XML, il incorpore également d'autres langages tels qu'OWL ou RDF permettant de créer des vocabulaires et des classifications d'objets.

Le web 4.0 transformera radicalement la forme de cette technologie. En effet, plusieurs experts prédisent l'optimisation des connexions sans fil et une intégration de ces connexions aux objets composant notre environnement **[web 8]**.

2.2 Définition de web 3.0

Le WEB 3.0 est le nom prévu pour la prochaine version de l'internet. C'est un concept émergeant qui s'articule autour du WEB sémantique et de l'intelligence artificielle qui pourrait être de type AMAS. Il est considéré comme l'internet 3ème génération, version évoluée de la version statique du WEB vers le WEB intelligent en passant par le WEB dynamique classique. Il aura pour mission de faire cohabiter le web est l'intelligence artificielle distribuée (IAD) en utilisant un ensemble, d'outils, protocoles, normes, Standards permettant à des machines et à des « agents web intelligents » d'effectuer, des raisonnements, des traitements...en ligne d'une manière coopérative et automatique sur des contenus **WEB[9].**

2.3. Evolution du web 3.0

Le WEB 3.0 est une évolution du WEB 2.0 et des propriétés qui le caractérisent. Ce qui veut dire que les aspects liés à l'interactivité et à la coopération feront toujours partie du WEB futur. C'est à dire aussi que dire que les utilisateurs ne seraient pas remplacés complètement par les machines, mais, auront, en plus des rôles joués dans la version actuelle du WEB, des rôles supplémentaires liés en particulier : au contrôle, à la supervision, à la configuration des applications et des agents. Ils seront en retrait dans certains cas pour laisser les machines effectuer les premiers traitements avant leur validation. Ainsi, les machines joueront un rôle important comme outils d'aide pour effectuer des traitements ou des prétraitements sur Internet quand l'intervention humaine n'est pas nécessaire. **[15]**

3. Le web sémantique

3.1. Définition

Le WEB sémantique est une notion très importante de l'Internet moderne, considéré comme intelligent où on ne se contente pas de stocker et diffuser des données, mais on s'intéresse à leur compréhension en effectuant des raisonnements sur leurs sens par des machines et agents logiciels. On peut le définir comme un ensemble de technologies permettant aux machines d'effectuer des traitements (Dans certains cas difficiles pour l'homme, par exemple : indexation, compréhension et recherche sémantiques) sur des données en s'appuyant sur les concepts comme, l'expression du sens, La représentation des connaissances, Les ontologies, Les agents, L'évolution de la connaissance **[15]**.

3.2 Acteurs

Un agent est une entité du système modélisé, situé dans un environnement, doté de capacités d'adaptation et d'autonomie lui permettant d'atteindre ses objectifs. Il existe plusieurs types d'agents. **[15]**

- Réactifs qui ne font que réagir aux stimuli qu'ils perçoivent d'une manière mécanique. Le comportement du système émerge des réactions simples des agents.
- Cognitifs disposant de capacités de raisonnement sur sa représentation du monde où ils évoluent. L'une des architectures les plus connues pour ce modèle est l'architecture BDI (Belif, Desir, Intentions) où les agents sont constitués principalement des caractéristiques suivantes :
  - ✓ Les croyances qui représentent la connaissance sur l'état de l'agent et de l'environnement où il évolue, c'est à dire ce que l'agent connaît sur lui-même et sur le monde où il évolue,
  - ✓ Les buts qui représentent la connaissance sur les motivations et les objectifs de l'agent,

- ✓ Les intentions qui représentent les informations sur les choix des plans que l'agent peut faire pour satisfaire des exécutions possibles.

Un système multi agents est un ensemble d'agents partenaires partageant des ressources et compétences complémentaires, similaires ou dissemblables et coopérant afin d'atteindre des objectifs partagés. Un Système Multi Agents Adaptatif (AMAS) est un système autonome qui doit faire face à des situations imprévues qui ne peuvent pas être résolues de manière algorithmique. Le comportement global (comportement émergent) d'un AMAS est le résultat de la coopération définissant l'organisation entre agents, ce qui revient à dire que, pour changer la fonction globale d'un AMAS, il suffit de changer l'organisation des agents le composant, dits agents AMAS. Les agents dans ce contexte, doivent faire face en permanence aux changements liés à l'environnement et aux situations non prévues. La communication au sein d'un AMAS entre agents se fait par l'intermédiaire d'un protocole d'interactions qui constitue un ensemble de règles de conduite que les agents doivent respecter entre eux afin de structurer leurs échanges. Et la mission de l'Internet futur, est de concilier ces deux notions dans le contexte du WEB, dans la perspective d'un WEB intelligent appelé WEB 3.0.

4. internet des objets

L'infographie proposée par Dassault remonte l'histoire des objets connectés jusqu'aux années 1800, période de l'invention du télégraphe, le premier appareil électronique de radiocommunication. En 1926, c'est la première fois où Nikola Tesla imagine un monde où les habitants seraient reliés par des installations de communication sans fil. A partir de 1989, Tim Berners-Lee pose les bases du World-Wide-Web, l'internet universel que nous connaissons aujourd'hui. **[Web10]**

Les premiers objets connectés n'apparaissent que dans les années 1990. Il s'agit de grille-pain, machines à café ou autres objets du quotidien. En 2000, le fabricant coréen LG est le premier industriel à parler sérieusement d'un appareil électroménager relié à internet, Les années 2000 verront les premières expérimentations d'appareils connectés à Internet. Ils l'utilisent notamment pour consulter des informations de matière automatique, notamment Ambiant Orb vers 2002. C'est bien avant qu'IPSO ne propose le concept d'adresses IP(2008) ou que l'organisation des Nations Unies ne mentionne l'internet des objets dans un rapport sur les télécommunications internationales. Depuis 2011, l'IPV6 offre de nouvelles possibilités pour les objets connectés qui disposent de nouvelles plages d'adresses IP disponibles et attribuables. En 2013, Intel lançait l'Internet of Things Solutions Group et nous créions ce site en janvier Dassault Systèmes recommande quelques articles sur l'internet des objets écrits au cours de l'année 2013, dont ceux des magazines en ligne américains Wired, Venture Beat ou CNBC, des sources qui nous inspirent également au quotidien.

A la fin de l'année 2012, il y avait environ 8,7 milliards d'objets connectés dans le monde. Cisco estime que ce nombre atteindra sans mal les 50 milliards d'objets connectés en 2020. **[Web10]**

4.1. Qu'est-ce qu'un objet ?

Un objet connecté est un objet physique équipé de capteurs ou d'une puce qui lui permettent de transcender son usage initial pour proposer de nouveaux services. Il s'agit d'un matériel électronique capable de communiquer avec un ordinateur, un smartphone ou une tablette via un réseau sans fil (Wi-Fi, Bluetooth, réseaux de téléphonie mobile, réseau radio à longue portée de type Sigfox ou LoRa, etc.), qui le relie à Internet ou à un réseau local.

On distingue communément deux grands groupes d'objets connectés :

- Les objets destinés à **la collecte et l'analyse de données**, dont la mission principale est de collecter et transmettre des informations.
- Les objets qui répondent à une logique de **contrôle-commande** et permettent de déclencher une action à distance **[Web11].**

4.2. Qui profit de l'internet des objets ?

L'internet des objets est une opportunité mais c'est également un énorme challenge à relever : le nombre d'identités qu'il faut gérer se développe très rapidement, bien plus vite que ce que la plupart des plateformes de gestion d'identités ne peuvent supporter. De nos jours, tout le monde parle de l'Internet des Objets, mais de quoi s'agit-il vraiment ? Au cas où vous vous posez la question, voici ce qu'il faut savoir : **l'Internet des Objets (IoT)** englobe tout ce qui se connecte à internet, incluant ainsi une variété toujours plus grande d'appareils tels que les Fitbit, les systèmes d'alarme pour domicile, les puces électroniques pour animal, les implants cardiaques, les codes RFID sur les articles d'inventaire, et les automobiles avec capteurs intégrés. D'après la société Gartner, l'Internet des Objets comptabilisera près de 26 milliards d'appareils d'ici 2020. **[Web12]**

4.3. Marché de l'internet des objets

Mc Kinsey voit l'Internet des objets représenter une opportunité business de 6200 milliards de dollars en 2025, quand Cisco table sur un marché de 14 400 milliards de dollars en 2022 et General Electric sur un chiffre de 15 000 milliards de dollars en 2034. Dans son étude "L'Internet des objets, premier marché à plus de 10 000 milliards de dollars, vraiment ?", Strategy Analytics se penche sur l'évolution du marché du machine-to-machine, précurseur de l'Internet des objets. "Avec une valeur globale de 3500 milliards de dollars pour l'ensemble du marché du numérique en 2016, il faut émettre des hypothèses héroïques pour parvenir à des estimations de l'Internet des objets à des milliers de milliards de dollars", se moque Andrew Brown, analyste chez Stratety Analytics. **[Web13].**

4.4. Comment ça marche ?

4.4.1. Que faut-t-il pour connecter les objets

Alors que les objets connectés se multiplient, les technologies de communication restent méconnues.

a. La communication courte-portée

Certaines technologies de communication sont à courte portée (faible distance entre émetteur et récepteur). L'échange de données peut d'abord se faire par **contact physique** entre l'émetteur et le récepteur, par exemple grâce à un port ethernet ou à un port USB. La **technologie** NFC (Near Field communication) consiste à transmettre des données sur des ondes haute fréquence (13,56 MHz). Ouvent, l'émetteur est **une puce RFID**, un composant passif (qui n'a pas besoin d'énergie externe pour fonctionner) qui comporte une antenne et une puce électronique associée à un identifiant unique. Lorsque le lecteur (pensez aux bornes de RATP) va demander à recevoir des données, il va transmettre le message de demande par des ondes qui vont également alimenter l'émetteur (pensez à votre carte NAVIGO) qui, en retour, va transmettre les informations désirées et stockées sur la puce électronique. La connexion à internet via un Hub. **[Web14]**

b. La connexion à internet via un Hub

Les technologies de courte portée vues précédemment sont adaptées dans le cas où la distance entre émetteur et récepteur est réduite (moins de 10cm). Pour connecter vos objets à l'internet dans le cas d'une distance de communication modérée, vous pouvez utiliser un Hub qui fera office d'interface entre l'internet virtuel et vos objets connectés. Les objets communiqueront au Hub, directement connecté à internet, via les technologies Bluetooth, Wi-Fi, Zigbee, Z-wave. **[Web14]**

c. Les réseaux cellulaires longues portées

Dans le cas d'objets qui ont besoin de pouvoir communiquer de longue distance, ou de zones difficilement accessibles, la solution de la communication courte portée et la solution de la communication par Hub sont inefficientes. Il est donc dans ce cas nécessaire de s'appuyer sur un réseau cellulaire qui permet une connexion en tout lieu couvert par les antennes.

Les objets connectés peuvent d'abord s'appuyer sur les réseaux cellulaires couverts par les opérateurs de téléphonie (2G, 3G, 4G, LTE) pour transmettre leurs données. **[Web14]**

d. La communication MESH

Certains objets connectés communiquent non pas vers un récepteur unique, ni vers un Hub, mais vers leurs pairs. Ainsi, chaque objet connecté est capable de recevoir et d'émettre

des données, ce qui forme un réseau maillé qui atteint le Hub, connecté à l'internet. L'avantage de ce type de communication est la redondance du message qui est transféré à chaque fois à tous les pairs de l'objet connecté qui émet le message **[Web14].**

4.4.2. L'identification des objets

L'identification par type ou par entité est une notion fondamentale de l'IoT. En général, les identifiants sont numériques. Par exemple, les produits de consommation ont généralement un code barre, les livres des ISBN, etc. Des objets isolés peuvent également avoir des numéros attribués : les puces RFID stockent des codes de produits électroniques grâce à des suites de 96-bits. Les adresses IP de nos ordinateurs sont un autre exemple d'identification. **[Web15]**

4.4.3 La technologie Near Field Technologie

La technologie NFC (Near Field communication) consiste à transmettre des données sur des ondes haute fréquence (13,56 MHz), permettant l'échange d'informations entre des périphériques jusqu'à une distance d'environ 10 cm. Les débits de transfert autorisés s'échelonnent jusqu'à 424 kbit/s. La technologie principalement est supportée par Sony, Philips, Nokia, Samsung et Panasonic. Aujourd'hui, la technologie NFC est utilisée dans les paiements carte bancaire par contact ou encore elle est présente sur les pass Navigo de notre métro Parisien. **[Web14]**

4.4.4 Les normes et les standards

A travers des normes communes que pourra émerger l'internet des objets. Les capteurs connectés nécessitent des processus de fond à l'image d'un nouveau protocole internet comme l'IPv6. De plus, de nouvelles normes de connectivité comme le processus IEEE 802.11ah permettront la connexion à basse puissance de nombres de périphérique tels les compteurs intelligents. Conscient de distribuer des données passivement, le consommateur entend bien voir son quotidien s'en améliorer. Ainsi, le suivi du sommeil (36%), les habitudes de conduite (35%), garder un œil sur ses proches vieillissants (34%) ou sur ses enfants (29%) font partie des attentes les plus prégnantes. Mais si les données d'observation forment un préalable nécessaire, c'est surtout la rétroaction de solutions a méliorative de ses habitudes et comportements qui constitue le véritable enjeu. **[Web16]**

4.5. Domaines d'utilisation

La technologie IOT révolutionne la vie quotidienne de tous les possesseurs de smart watches, de réfrigérateurs intelligents, de voitures connectées… mais aussi le monde de l'industrie et de la santé. L'internet des objets permet de mieux gérer les coûts des entreprises et permet de fournir des résultats de meilleurs de qualités **[Web17].**

5. Modèles de référence

L'Internet des Objets demande un modèle de référence qui permettrait de décrire la manière avec laquelle ces systèmes, ces réseaux et ces applications interagissent entre eux. En effet, un tel modèle aurait les avantages : [**Web18**]

- **De simplifier** la compréhension de systèmes complexes découpés en parties plus compréhensibles
- **De clarifier** en fournissant des informations supplémentaires identifiant les niveaux de l'IoT et fournissant une terminologie commune
- **D'identifier** où des types spécifiques de traitement sont optimisés dans les différentes parties du système
- **De standardiser** pour créer les conditions d'une interopérabilité entre des produits IoT de différents fabriquant
- **D'organiser** rend l'IoT plus accessible et moins conceptuel

Cisco Systems propose un modèle en sept couches qui découpe les opérations IoT en sept niveaux distincts chacun correspond à une fonction dans le processus de traitement. Il ne faut pas voir ce modèle comme étant strictement défini quant aux composants ou aux endroits. Par exemple chaque fonction peut être combinée dans une seule armoire dans un centre de données ou être distribuée sur différents périphériques répartis dans le monde. Le modèle explique comment la tâche exécutée au niveau de chaque couche peut être maintenue de manière simple, avec haute disponibilité et support. Enfin, le modèle définit les conditions pour disposer d'un système IoT complet voir figure 1.10.[**Web18**]

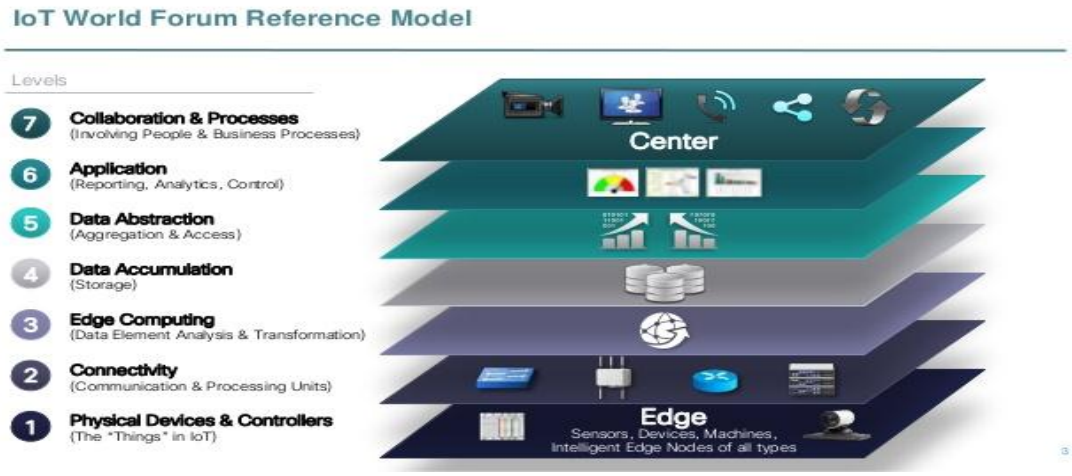

Figure3.1 model de référence IOT **[Web18]**

Le modèle de référence IoT en sept couches décrit deux flux :

- Dans un modèle de contrôle, un flux descendant, de la couche 7 à la couche 1
- Dans un modèle de surveillance, un flux montant, de la couche 1 à la couche 7 (Voir figure 2)

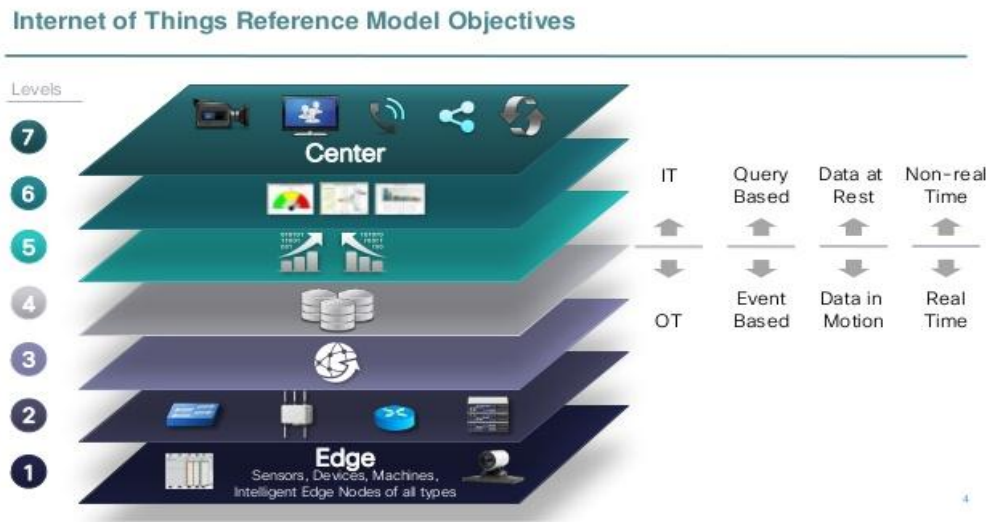

Figure 3.2 Objectives de model de référence IOT **[WEB18]**

6. Protocoles IOT

Les futurs standards de communication connectant le vaste ensemble de l'internet des objets ne sont pas encore définis et font l'objet d'une âpre bataille, aux enjeux tant techniques et économiques que politiques. «La communication sur internet se base sur une couche IP (protocole internet) qui harmonise à un certain niveau tous les standards existants, mais l'internet des objets possède des protocoles propres à certains métiers et usages, loin de garantir une communication unifiée et sécurisée entre des équipements hétérogènes. [**Web19**]

6.1. Les Protocoles d'accès

Les performances des systèmes de radiocommunications sont fortement liées aux choix techniques qui permettent à des utilisateurs multiples (multi user) d'accéder à un canal de transmission. Ce cours aborde les mécanismes d'allocations des ressources physiques que ces dernières soient des fréquences, des times slots ou des codes d'étalement.

Cet ensemble de mécanismes constitue une sous couche, appelée Medium Acces Control (MAC), de la couche "Liaison de Données" définie par l'ISO

On rappelle qu'il s'agit en fait d'une sous couche de la couche liaison dont le but principal est d'attribuer un canal à accès multiples à plusieurs utilisateurs. Ce type d'allocation revêt une importance capitale dans les systèmes radio dit point multipoints pour lesquels une station centrale, station de base d'un système cellulaire terrestre ou station terrienne d'un système satellitaire, est reliée à un grand nombre d'utilisateurs équipés de terminaux. **[18]**

Deux problèmes difficiles doivent alors être résolus. **[16]**

1. Le premier concerne les mécanismes d'allocation qu'il faut mettre en œuvre dans le nœud d'accès pour faire la correspondance entre les demandes des ressources des différentes communications en cours et les ressources disponibles.
2. L'autre problème concerne les différentes solutions qu'a un terminal, qui ne dispose d'aucune ressource, de faire savoir à sa station de base qu'il a besoin de ressources pour faire "passer" une communication.

6.1.1 Protocoles M2M

La communication machine a machine est l'association des technologies de l'information avec des objets dits intelligents et communicants et cela dans le but de fournir à ces derniers les moyens d'interagir sans intervention humaine avec le système d'information. Ce dernier peut appartenir indifféremment à une organisation ou à une entreprise. Les principaux protocoles et les différentes solutions de réseaux qui permettent de mettre en œuvre une application M2M sont :

a. Z-Wave :

Est une technologie radio opérant sur les bandes de fréquence 868/915, il est aussi utilisé dans le domaine de la domotique. C'est une norme assez puissante pour l'automatisation de structures complexes. Avec une bande passante sur 8 bits qui est plus que suffisant pour allumer et éteindre des lampes. Le Z-Wave est un protocole facile à intégrer et est supporté par la plupart des dispositifs domotiques récents.

b. Zigbee:

Avec une bande passante de 256 bits qui permet l'utilisation de dispositifs de communication plus complexes permettant la transaction de plus grandes quantités d'informations tels que des dispositifs médicales, domotiques etc...les équipements Zigbee sont plus chers que les équipements Z-Wave

c. Bluetooth Low Energy ou BLE:

Apparu vers 2011 sous l'appellation **Bluetooth 4.0**, le Bluetooth Low Energy consomme moins d'énergie que le Bluetooth traditionnel ce qui est très bénéfique pour les applications M2M qui ne demandent qu'à échanger périodiquement de petites quantités de données. Avec une faible consommation d'énergie, les équipements peuvent fonctionner sur une petite batterie sur une longue durée.

d. 6LoWPAN:

Opérant sur les fréquences 868 MHz, 900 MHz et 2.4 GHz comme le Wifi ou le Bluetooth, le 6LoWPAN ou ***Low Power Wireless Personal Area Network*** permet la

transmission de données d'appareils à faible puissance via le protocole IPv6.Concurrent du Zigbee, le 6LoWPAN peut faire communiquer des appareils via des réseaux IP comme le Wifi.

e. NFC :

Le NFC ou Near Field Communication est une technologie de communication sans fil à haute fréquence et à courte portée qui permet l'échange de données entre des périphériques sur environ 10 centimètres. Le NFC est principalement destiné à une utilisation dans les téléphones mobiles ou tablettes tactiles. **[Web20]**

6.1.2. Protocole LAN

Les protocoles LAN sans fil se distinguent par leur capacité à fournir efficacement des données sur de courtes distances, comme quelques centaines de mètres, à travers différents médiums, tels que les câbles de cuivre. Différents protocoles existent à des fins différentes et existant dans différentes « couches » de l '« Open SystemsInterconnect », ou OSI.

Typiquement, lorsque vous utilisez le mot « LAN » pour décrire un protocole, l'intention est de décrire le niveau inférieur, ou physique. Certains des protocoles LAN les plus courantes sont "Ethernet", "Token Ring" et "Distributed Data Interface Fibre", ou "FDDI".

a. Ethernet

Est de loin le type le plus commun de protocole de réseau local. On le trouve dans les maisons et bureaux à travers le monde et est reconnaissable par son support commun "CAT5" de cuivre du câble.

b. Token Ring

Est une technologie de réseau local ancien qui n'est pas répandue plus. Le principe de base de " Token Ring " est un "jeton" est passé d'un système à, ou via un concentrateur, et seul le destinataire lit le jeton.

c. FDDI

Définit la manière dont le trafic LAN est transmis sur un câblage de la fibre. Câbles à fibres optiques est utilisé lorsque les distances plus longues, généralement entre les étages ou bâtiments, sont nécessaires, ou lorsque la sécurité accrue est nécessaire.

6.1.3. Protocole WAN :

Protocoles WAN se distinguent par leur capacité à fournir efficacement des données sur de plus longues distances, comme des centaines de miles. Ceci est généralement nécessaire pour combler des données entre plusieurs réseaux locaux. L'Internet est le plus grand réseau

étendu du monde. Routeurs, modems et autres dispositifs WAN sont utilisés pour transmettre des données sur divers supports, câblage fibre couramment. Certains des protocoles WAN les plus couramment utilisés aujourd'hui sont « Frame Relay », « X25 », « Integrated Services Digital Network », ou « RNIS » et « protocole Point-to -Point " ou " PPP **[Web21]**

a. Frame Relay " et " X.25 "

Ils sont similaires en ce sens qu'ils sont tous deux des technologies de commutation par paquets pour l'envoi de données sur de grandes distances. " Frame Relay " est plus récent et plus rapide, alors que " X.25 " fournit des données plus fiables.

b. "PPP"

C'est un protocole qui est utilisé pour transmettre des données à d'autres protocoles sur les médiums qu'ils ne soutiendraient pas normalement, comme l'envoi de la "protocole Internet", ou IP, sur des lignes série

c. «RNIS»

C'est une méthode de combiner plusieurs lignes à distance sur un réseau téléphonique public en un seul flux de données. **[Web22]**

6.2. Protocoles applicatifs

Toutes les applications Internet complexes utilisent des protocoles de transmission spécifiques, qui viennent se greffer sur les protocoles TCP/IP. Ces protocoles sont définis en fonction de leurs applications respectives. Le protocole assurant la transmission des documents HTML sur le World Wide Web est évidemment d'une importance primordiale. Ce protocole est HTTP, en outre, les navigateurs Web disposent d'outils leur permettant de réceptionner et de traiter des informations transmises à l'aide d'autres protocoles Internet majeurs. **[Web23]**

6.2.1. Quelques protocoles applicatifs

a. IPv4

C'est un protocole « routable », de la couche réseau du modèle OSI (couche 3), entendez par là qu'il définit principalement un système d'adressage permettant de router des paquets**. [Web17]**

b. Le protocole IPv6

Il répond raisonnablement aux objectifs édictés. Il maintient les meilleures fonctions d'IPv4, en écarte ou minimise les mauvaises, et en ajoute de nouvelles quand elles sont nécessaires. En général, IPv6 n'est pas compatible avec IPv4, mais est compatible avec tous les autres protocoles Internet, dont TCP, UDP, ICMP, IGMP, OSPF, BGP et DNS ; quelque

fois, de légères modifications sont requises (notamment pour fonctionner avec de longues adresses). **[Web25]**

c. Le protocole 6LoWPAN

Il a été développé pour définir l'adaptation d'IPv6, ainsi que la manière de transporter les datagrammes IP sur des liaisons IEEE 802.15.4 et d'exécuter les fonctions de configurations nécessaires pour former et maintenir un sous-réseau IPv6. **[Web26]**

d. TCP/UDP

- **UDP** est un protocole orienté "non connexion". Pour faire simple, lorsqu'une machine A envoie des paquets à destination d'une machine B, ce flux est unidirectionnel.
- Contrairement à l'UDP**, le TCP** est orienté "connexion". Lorsqu'une machine A envoie des données vers une machine B, la machine B est prévenue de l'arrivée des données, et témoigne de la bonne réception de ces données par un accusé de réception. **[Web27]**

e. HTTP

Il sert à transmettre des documents au format (HTM, STM, SSI, etc.)HTML ou autres sources d'informations demandés au serveur par le navigateur, contrairement au protocole FTP entre autres, du cas présent la connexion est ici temporaire. Elle s'interrompt dès la transmission effectuée, sans demander confirmation à l'utilisateur. En principe, l'utilisateur d'un navigateur Web ne peut pas intervenir directement sur la connexion HTTP que le logiciel est en train d'établir. Le numéro du port donnant accès aux serveurs HTTP est le 80 (en générale). **[Web28]**

f. UPNP

UPnP, comme son nom l'indique, est dérivé de PnP (Plug aNd Play), qui est une technologie qui permet de faciliter l'installation, la configuration et l'ajout de périphériques informatiques à un micro-ordinateur. Universal Plug And Play (UPnP) étend cette simplicité en incluant l'ensemble du réseau informatique, permettant la découverte et le contrôle des périphériques, y compris les dispositifs et services en réseau, tels que les imprimantes connectées au réseau ou les équipements électroniques. **[Web29]**

g. COAP

CoAP (Constrained Application Protocol) est un protocole de transfert Web optimisé pour les périphériques et réseaux contraints utilisés dans les réseaux de capteurs sans fil pour former l'Internet des objets. Basé sur le style architectural REST, il permet de manipuler au travers d'un modèle d'interaction client-serveur les ressources des objets communicants et capteurs identifiées par des URI en s'appuyant sur l'échange de requêtes-réponses et méthodes similaires au protocole HTTP.

h. MQTT

MQ Telemetry Transport (MQTT) est un protocole open source qui a été développé et optimisé pour les appareils limités et à faible bande passante, à latence élevée, ou pour les réseaux non fiables. Il s'agit d'un transport de messages de publication/d'abonnement qui est extrêmement léger et idéal pour le raccordement à des réseaux de petits appareils avec une bande passante minimale. **[Web30]**

i. XMPP

XMPP (abréviation d'Extensible Messaging and Presence Protocol, anciennement connu sous le nom de **Jabber**) est un protocole de messagerie instantanée. Ce protocole sert de support à *Google Talk*, *Google Wave*, *Gizmo5*, *IBM Lotus Notes*, *Facebook*, etc., qui sont des ensembles de protocoles. Il a été créé par Jeremie Miller qui l'a lancé sous le nom de *jabberd* en 1998. **[Web31]**

7. Architectures IOT

Une architecture IOT peut être vue comme suit **[Web32]**

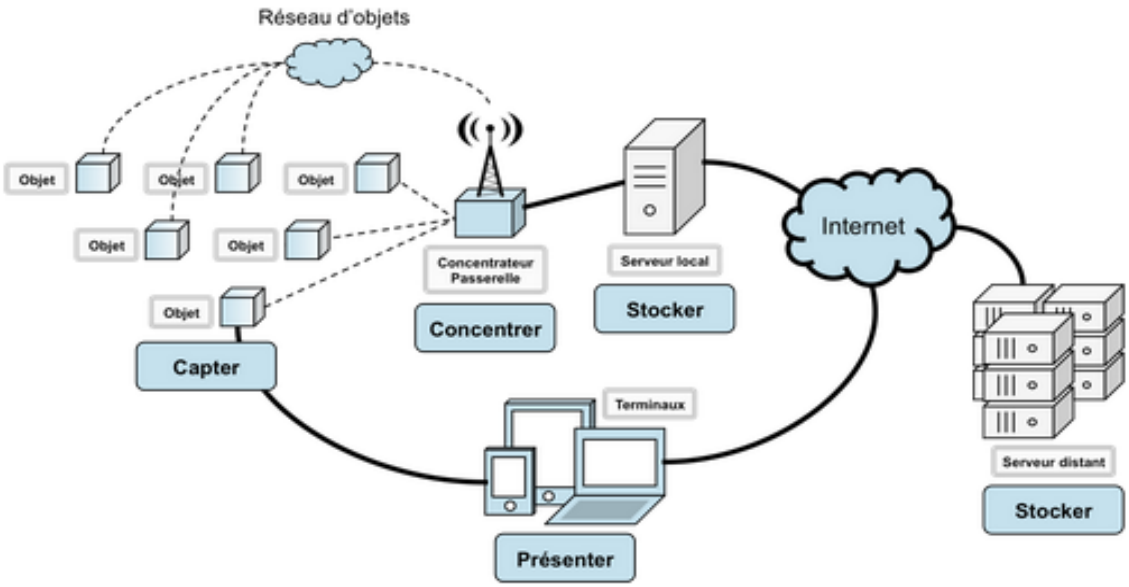

Figure 3.3: Architecture IOT **[Web32]**

- **Une fonction Capter** désigne l'action de transformer une grandeur physique analogique en un signal numérique.
- **Une fonction Concentrer** permet d'interfacer un réseau spécialisé d'objet à un réseau IP standard (ex. Wifi) ou des dispositifs grand public.

- **Une fonction Stocker** qualifie le fait d'agréger des données brutes, produites en temps réel, méta taguées, arrivant de façon non prédictible.

- Une fonction **Présenter** indique la capacité de restituer les informations de façon compréhensible par l'Homme, tout en lui offrant un moyen d'agir et/ou d'interagir.

Deux autres processus n'apparaissent pas sur le schéma, car ils sont à la fois transverses et omniprésents :

- Le **traitement des données** est un processus qui peut intervenir à tous les niveaux de la chaîne, depuis la capture de l'information jusqu'à sa restitution. Une stratégie pertinente, et commune quand on parle d'Internet des objets, consiste à stocker l'information dans sa forme intégrale.
- **La transmission des données** est un processus qui intervient à tous les niveaux de la chaîne. Deux réseaux, supports des transmissions, cohabitent généralement :

  - **Réseau local de concentration** On utilise alors des technologies comme ANT, ZigBee, Z-wave, NFC ou Bluetooth LE.
  - **Réseau WAN** permettant d'interconnecter les réseaux spécialisés et de les interfacer avec des fermes de serveur. On utilise alors Wifi, les réseaux cellulaires (GSM, UMTS, LTE) ou encore les connexions physiques standard (Ethernet, fibre optique). Ces réseaux sont généralement connectés à Internet. **[Web32]**

Conclusion

Le marché des nouvelles technologies est marqué par une nouvelle discipline : les « objets connectés ». Ces concentrés technologiques sont des objets qui utilisent Internet pour améliorer leur fonctionnement, en étant souvent l'évolution technique d'un objet déjà existant.

Le premier fut la domotique : la technologie du domicile. C'est l'un des domaines les plus importants des objets connectés, où ils jouent un rôle de contrôle à différentes échelles. Ainsi, d'une application de téléphone, il est désormais possible d'allumer ses lumières, d'ouvrir le portail ou de prévoir le chauffage pour son retour de vacances. Le quotidien est alors facilité.

Les objets connectés peuvent aussi avoir des tâches plus importantes et plus complexes, c'est le cas notamment dans le domaine de la santé, où ils deviennent des capteurs précis et facile à transporter. Ces facultés sont très importantes dans le cadre de certaines maladies : épilepsie, diabète, qui nécessite un contrôle permanent, d'un côté de l'activité corporelle, de l'autre de la glycémie.

Le troisième et dernier domaine est celui du quotidien. Nous y retrouvons le sport : les trackers d'activité au poignet sont de plus en plus nombreux, permettant aux sportifs d'analyser leurs performances. Les loisirs ne sont pas oubliés, avec les Smart watchs et les lunettes connectées, telles que les GoogleGlass, qui ouvrent un nouvel horizon à l'implication de la technologie dans la vie de tous les jours. Limiter les objets connectés aux hommes serait réducteurs, et les projets les plus ambitieux souhaitent aménager les villes afin qu'elles possèdent elles-mêmes une connectivité.

Toute cette première partie a permis d'émettre une affirmation : les objets connectés sont une source de progrès. Ils ouvrent des possibilités et faciliteront la vie quotidienne d'ici quelques années.

Cependant, l'avenir n'est pas tout rose, et il est nécessaire de rester pragmatique à ce sujet. Bien qu'ils soient intéressants, les objets connectés amènent tout un lot de problèmes qu'il faudra prendre en compte tant ils sont importants : c'est ce que nous avons vu durant la seconde partie de notre recherche.

Le premier risque provient de l'homme lui-même. Le fait de s'entourer de technologie informatise la vie : le piratage peut alors avoir des conséquences désastreuses. Un cracker pourrait facilement s'infiltrer dans le réseau de votre maison, et la contrôler comme il l'entend : bien que cela puisse faire sourire en premier lieu, au niveau de la sécurité c'est bien différent. Le cracker peut également avoir une surveillance de la vie personnelle, à l'aide par exemple des webcams ou caméras de smartphones. C'est encore pire pour la santé : la vie des malades entre très rapidement en jeu.

Le second risque est de baser toute la vie aveuglément sur les objets connectés, ce qui, en cas de panne, serait très dangereux : la vie s'arrêterait tout simplement, à petite ou à grande échelle. La panne a des causes souvent imprévisibles, sauf avec une vérification très précise quotidiennement : cela est impossible. Il est aussi nécessaire de prévoir le besoin énergétique pour que tous ces objets puissent fonctionner : ils se compteront en dizaines de milliards d'ici 2030, et représenteront donc une part non négligeable de la consommation électrique mondiale.

Le dernier risque est celui de la dérive. A travers toutes nos recherches, nous avons pu voir un nombre gigantesque d'objets, et nous avons parlé des plus intéressants. Il est cependant essentiel de préciser que ces exemples ne sont qu'une infime minorité de l'immensité des objets connectés, et un constat se fait rapidement : sont-ils réellement tous utiles ? Ne pourrait-on tout simplement pas s'en passer pour la plupart ? De plus, un développement de maladies déjà existantes (par exemple les TOC) apparait avec les objets connectés, et le stress lié au besoin de contrôle s'aggrave de plus en plus. Sont-ils, compte tenu de tous ces défauts, réellement intéressants ?

Le futur répondra certainement à ces interrogations : le progrès est en marche, et il est inarrêtable. Nous finirons en citant Albert Einstein à propos de la technologie : « *Il est hélas devenu évident aujourd'hui que notre technologie a dépassé notre humanité.* ».

Pour conclure, l'IoT incarne la prochaine évolution de l'Internet. Sachant que l'être humain progresse et évolue en transformant les données en informations, en connaissances et en savoir, l'IoT a le potentiel d'améliorer le monde tel que nous le connaissons. La rapidité à laquelle nous y parviendrons ne dépend que de nous.

Nous sommes aux débuts de l'Internet des objets car il existe encore beaucoup de questions ouvertes. L'écosystème doit se structurer. Si de nouveaux acteurs vont apparaître pour simplifier l'usage, d'autres disparaîtront faute d'avoir pu évoluer et adopter les standards. Il est possible d'esquisser un parallèle avec l'apparition du Web qui, au milieu des années 90 a révolutionné le réseau Internet. Les premières pages étaient sommaires, composées manuellement, puis sont venues des sociétés comme Facebook qui ont su démocratiser leur usage en proposant de nouveaux services innovants.

D'un point de vue protocolaire, si de nombreux efforts de standardisation ont été accomplis par différents organismes de normalisation, il reste encore beaucoup de chemin à parcourir pour pouvoir exploiter ces réseaux à grande échelle.

Les protocoles et les architectures de l'Internet devront encore évoluer pour prendre en compte les diversités d'accès et de contenus, le très grand nombre d'équipements communicants et la garantie de confidentialité des données. En plus de ces enjeux technologiques, le principal défi industriel sera de prendre en compte des cycles de vie différents de l'informatique traditionnelle.

Chapitre 3 : Internet des objets IOT

L’internet des objets donne une idée des possibilités offertes par un certain nombre de technologies existantes et futures qui, ensemble, pourraient, dans les prochaines années, modifier en profondeur le mode de fonctionnement de nos sociétés. C’est une évolution majeure de nos systèmes d’information et de communication qu’entrainera l’internet des objets. Mais l'acceptation de l'IdO par la société sera fortement liée au respect de la vie privée et à la protection des données personnelles. Il est très probable que dans les années à venir, nous soyons confrontés aux problèmes d’interopérabilités, d’éthiques et de sécurités.

En adoptant une approche volontaire, les acteurs du développement de l’IdO peuvent jouer un rôle de premier plan pour définir les modalités de fonctionnement de l'IdO et retirer les bénéfices qui en découlent en termes de croissance économique et de bien-être individuel, faisant ainsi de l’internet des objets un internet des objets pour les individus.

Chapitre 4

Chapitre 4 : Sécurité d'internet des objets IOT

1. Introduction

L'IoT (Internet of Things ou Internet des objets) promet d'être l'une des tendances majeures de notre ère numérique.

Avec un défi majeur à relever, celui d'assurer la confidentialité et la sécurité des données échangées. Si l'on estime qu'il existe déjà des problèmes de sécurité avec un milliard de smartphones pourtant placés sous contrôle humain, quelle sera la situation avec mille milliards d'objets autonomes connectés, recueillant des informations sur notre santé, notre conduite automobile ou les paramètres de notre habitat ?

La connaissance humaine est pareille à une sphère qui grossirait sans cesse : à mesure qu'augmente son volume, le nombre de ses points de contact avec l'inconnu grandit. Ainsi, l'importance et la diversité attendue des informations générées, distribuées et traitées au sein de l'IoT ne nous permettent pas, pour le moment, d'appréhender complétement l'ensemble des menaces et vecteurs d'attaque qu'il va falloir affronter. Face à cet obstacle, la plate-forme architecturale des objets connectés au sein de l'IoT va jouer un rôle fondamental.

« The National Intelligence Council (NIC)» américain considère que les avancées technologiques combinées à une forte demande des marchés encourageraient une adoption et un déploiement à large échelle de l'IdO. Néanmoins, la plus grande crainte est que les objets du quotidien deviennent des risques potentiels d'attaque de sécurité. Pire encore, la pénétration à large échelle de l'IdO diffuserait ces menaces d'une façon beaucoup plus large que l'Internet d'aujourd'hui **[19]**.

En effet, l'ubiquité de l'IdO amplifiera les menaces classiques de sécurité qui pèsent sur les données et les réseaux. Mais en plus, le rapprochement du monde physique et du monde virtuel à travers l'IdO ouvre la voie à de nouvelles menaces qui pèseront directement sur l'intégrité des objets eux-mêmes, les infrastructures et processus (monde physique), et la vie privée des personnes.

2. Sécurité des systèmes informatiques

Avant de préciser la notion de sécurité des systèmes informatiques, il convient de définir au préalable ce que nous entendons par le terme système. Nous précisons ensuite la terminologie et les concepts fondamentaux liés à la sûreté de fonctionnement dans laquelle s'inscrit plus particulièrement la sécurité des systèmes. Finalement, nous rappelons la définition de la sécurité des systèmes informatiques telle qu'elle a été initialement décrite dans les Information Technology Security Evaluation Criteria. **[20]** puis, par la suite, reprise dans les Critères Communs **[21]**, c'est-à-dire la sécurité comme la combinaison de trois propriétés : la confidentialité, l'intégrité et la disponibilité de l'information.

2.1. Concepts et terminologie des systèmes [22]

2.1.1. Un système

Est une entité qui interagit avec d'autres entités, donc d'autres systèmes, y compris le matériel, le logiciel, les humains et le monde physique avec ses phénomènes naturels.

Ces autres systèmes constituent l'environnement du système considéré. La frontière du système est la limite commune entre le système et son environnement.

2.1.2. La fonction d'un système

Est ce à quoi il est destiné. Elle est décrite par la spécification fonctionnelle. Le comportement d'un système est ce que le système fait pour accomplir sa fonction est décrit par une séquence d'états. L'ensemble des états de traitement, de communication, de mémorisation, d'interconnexion et des conditions physiques constituent son état total.

2.1.3. La structure d'un système

Est ce qui lui permet de générer son comportement. D'un point de vue structurel, un système est constitué d'un ensemble de composants interconnectés en vue d'interagir. Un composant est un autre système, etc. La décomposition s'arrête lorsqu'un système est considéré comme étant un système atomique : aucune décomposition ultérieure n'est envisagée ou n'est envisageable, soit par nature, soit parce que dénuée d'intérêt.

2.1.4. Le service délivré par un système

Dans son rôle de fournisseur, est son comportement tel que perçu par ses utilisateurs ; un utilisateur est un autre système qui reçoit un service du fournisseur.

Un système peut être, séquentiellement ou simultanément, fournisseur et utilisateur d'un autre système, c'est-à-dire délivrer un service à cet autre système et en recevoir. La partie de la frontière du système où ont lieu les interactions avec ses utilisateurs est l'interface de service.

La partie de l'état total du fournisseur qui est perceptible à l'interface de service est son état externe ; le reste est son état interne.

Il est à noter qu'un système accomplit généralement plusieurs fonctions et délivre plusieurs services. Les concepts de fonction et de service peuvent donc être vus comme constitués de fonctions élémentaires et de services élémentaires.

2.2. Sûreté de fonctionnement

La sûreté de fonctionnement fournit un cadre conceptuel intéressant pour situer la sécurité par rapport à d'autres propriétés des systèmes informatiques. En effet, depuis de nombreuses années, les spécialistes de la sûreté de fonctionnement ont développé une terminologie et des méthodes dont l'application à la sécurité peut être enrichissante. Nous rappelons, dans cette sous-section, les définitions et les concepts de base de la sûreté de fonctionnement directement adaptés de **[23]** et de **[22]**.

La sûreté de fonctionnement d'un système est la propriété qui permet à ses utilisateurs de placer une confiance justifiée dans le service que le système leur délivre.

Cette propriété englobe trois notions différentes (Figure 1.13) : ses attributs, les propriétés complémentaires qui la caractérisent ; ses entraves, les circonstances indésirables – mais non inattendues– qui sont causes ou résultats de la non-sûreté de fonctionnement ; ses moyens, les méthodes et techniques qui cherchent à rendre un système capable d'accomplir correctement sa fonction et à donner confiance dans cette aptitude.

2.2.1. Attributs de la sûreté de fonctionnement

Les attributs de la sûreté de fonctionnement sont définis par diverses propriétés dont l'importance relative dépend de l'application et de l'environnement auxquels est destiné le système informatique considéré :

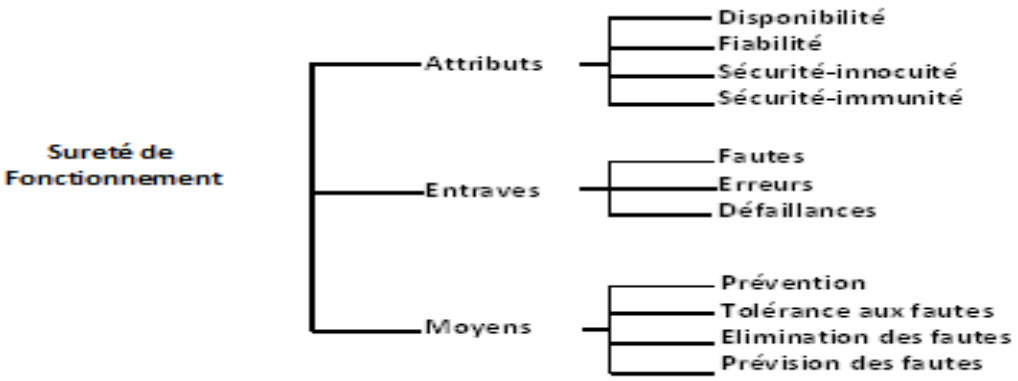

Figure 4.1 – Arbre de la sûreté de fonctionnement

2.2.2. Entraves à la sûreté de fonctionnement

Un service correct est délivré par un système lorsqu'il accomplit sa fonction. Une défaillance du service, souvent simplement dénommée défaillance, est un événement qui survient lorsque le service dévie de l'accomplissement de la fonction du système. Le service délivré étant une séquence d'états externes, une défaillance du service signifie qu'au moins un état externe dévie du service correct. La partie de l'état du système qui dévie du fonctionnement correct, autrement dit qui est anormale ou incorrecte, est une erreur ; une faute est la cause adjugée ou supposée d'une erreur, et peut être interne ou externe au système.

La relation de causalité entre fautes, erreurs et défaillances peut être exprimée comme suit.

- **Une faute activée** produit une erreur qui peut se propager dans un composant ou d'un composant à un autre, et est susceptible de provoquer une défaillance.
- **La défaillance d'un composant** cause une faute permanente ou temporaire interne pour le système qui le contient, tandis que la défaillance d'un système cause une faute permanente ou temporaire externe pour les systèmes avec lesquels il interagit. Ce processus de propagation est illustré sur **la figure 4.2**

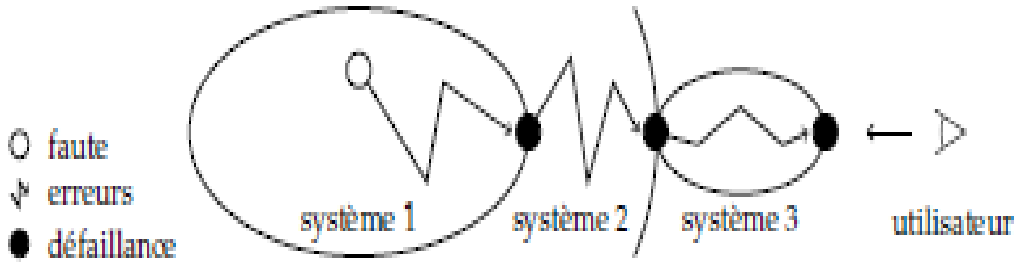

Figure 4.2 – propagation des erreurs dans un système de systèmes [24]

Deux principales classes de fautes sont à considérer dès lors que nous nous intéressons aux fautes malveillantes : les logiques malignes et les intrusions. Nous reprenons les définitions de ces termes telles qu'elles ont été introduites dans le projet MAFTIA (Malicious- and Accidental- Fault Tolérance for Internet Applications) **[24]**. **Les logiques malignes** sont des parties du système conçues pour provoquer des dégâts (les bombes logiques, les virus, etc.) ou pour faciliter des intrusions futures au travers de vulnérabilités créées volontairement telles que les portes dérobées. Les logiques malignes peuvent être introduites dès la création du système par un concepteur malveillant, ou en phase opérationnelle par l'installation d'un composant logiciel ou matériel contenant un cheval de Troie ou par une intrusion. **La définition d'une intrusion** est étroitement liée aux notions d'attaque et de vulnérabilité. Une attaque est une faute d'interaction malveillante visant à violer une ou plusieurs propriétés de sécurité. C'est une faute externe créée avec l'intention de nuire. Une vulnérabilité est une faute accidentelle ou intentionnelle (avec ou sans volonté de nuire) dans la spécification des besoins, la spécification fonctionnelle, la conception ou la configuration du système, ou dans la façon selon laquelle il est utilisé. La vulnérabilité peut être exploitée pour créer une intrusion. Une intrusion est une faute malveillante interne, mais d'origine

externe, résultant d'une attaque qui a réussi à exploiter une vulnérabilité. Elle est susceptible de produire des erreurs pouvant provoquer une défaillance vis-à-vis de la sécurité, c'est-à-dire une violation de la politique de sécurité du système.

2.2.3. Moyens pour la sûreté de fonctionnement

Le développement d'un système sûr de fonctionnement passe par l'utilisation combinée d'un ensemble de méthodes qui sont réparties en quatre classes de moyens :

- **La prévention de fautes** : comment empêcher, par construction, l'occurrence ou l'introduction de fautes ; elle est principalement obtenue par des méthodes de spécification et de développement relevant de l'ingénierie des systèmes ;

- **La tolérance aux fautes** : comment fournir un service à même de remplir la fonction du système en dépit des fautes ; elle est mise en œuvre par la détection d'erreurs et le rétablissement du système ;

- **L'élimination des fautes** : comment réduire le nombre et la sévérité des fautes ; elle peut être réalisée pendant la phase de développement d'un système par vérification, diagnostic et correction, ou pendant sa phase opérationnelle par maintenance ;

- **La prévision des fautes** : comment estimer la présence, la création et les conséquences des fautes ; elle est effectuée par évaluation du comportement du système par rapport à l'occurrence des fautes, à leur activation et à leurs conséquences.

2.3. Les fonctions principales de la sécurité informatique

Actuellement, l'information dans l'entreprise est une importance primordiale, ce qui rend sa protection est une fonction préliminaire. La sécurité des systèmes informatiques est devenue un défi majeur dont l'objectif est d'assurer la disponibilité des services, la confidentialité et l'intégrité des données et des échanges.

La sécurité informatique dépend de :

- **La confidentialité :** Assurer que l'information n'est pas mise à disposition à des personnes, des entités ou des processus non autorisés.
- **L'intégrité :** assurer l'exactitude et la complétude de l'information pour éviter la modification non autorisée de données.
- **La disponibilité :** assurer que l'information est accessible et utilisable sur demande par une entité autorisée.
- **Le non répudiation** : assurer qu'une action d'une entité peut être liée uniquement à son initiateur.

3. Les attaques

Sur internet des attaques ont lieu en permanence, à raison de plusieurs attaques par minute sur chaque machine connectée. Ces attaques sont pour la plupart lancées automatiquement à partir de machines infectées (par des virus, chevaux de Troie, vers, etc.), à l'insu de leur propriétaire. Plus rarement il s'agit de l'action de pirates informatiques **[web33]**.

3.1. Définition :

Une « attaque » est l'exploitation d'une faille d'un système informatique (système d'exploitation, logiciel ou bien même de l'utilisateur) à des fins non connues par l'exploitant du système et généralement préjudiciables.

Tout acte sur un système dont l'intention est de nuire au moins à l'une des propriétés de sécurité est qualifié de malveillant et constitue, de ce fait, une attaque sur ce système. Nous trouvons dans la littérature des manières différentes de classer les attaques.

Certaines taxonomies les organisent en fonction d'un unique critère. Parmi ces critères, les plus récurrents sont :

- **La cause de l'attaque** (utilisateur interne ou externe, intrus, etc.) ; **[25]**
- **Le mode ou le type de l'attaque** (virus, ver, écoute passive, déguisement, etc.) ; **[26]**
- **Le résultat de l'attaque** (divulgation, perturbation, etc.) **[27]**;
- **La vulnérabilité exploitée par l'attaque** ; plusieurs aspects peuvent alors être considérés : la phase de création ou d'occurrence de la vulnérabilité (lors de la conception, du développement, de l'exploitation du système, etc.), la nature de la vulnérabilité (concurrence, défaut de validation, etc.). **[28]**

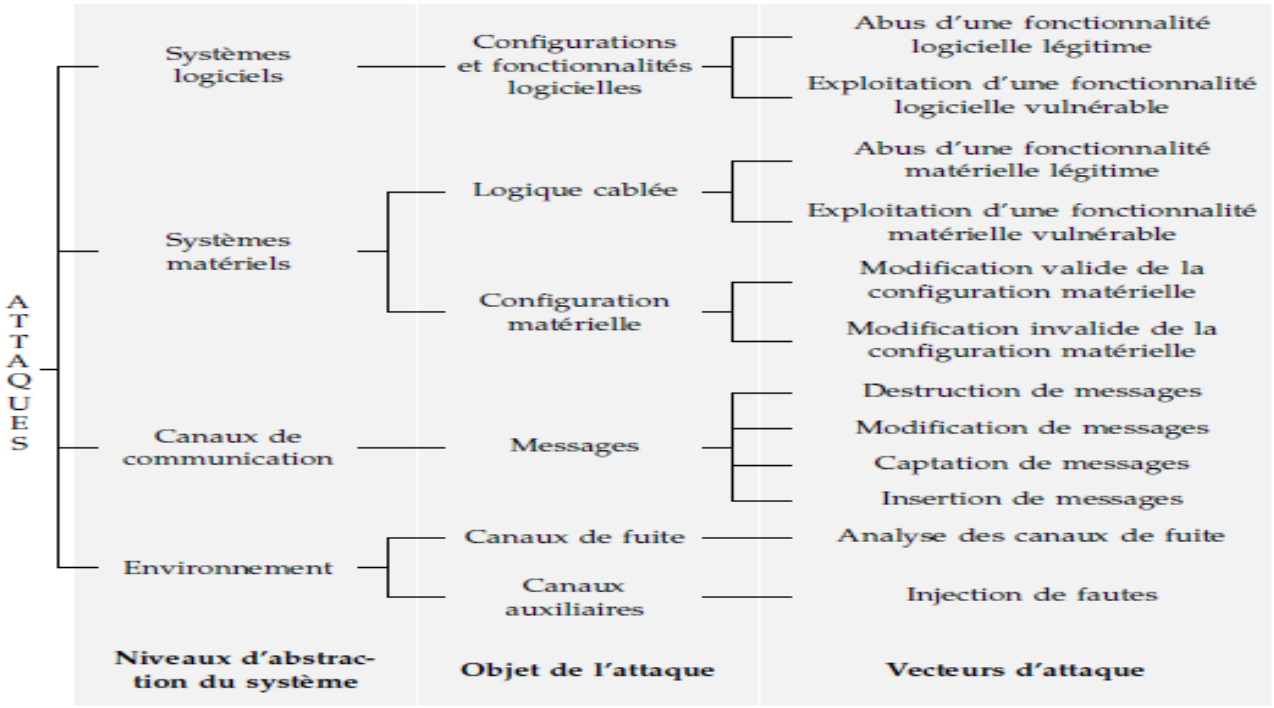

Figure 4.3 – Classification des attaques sur les systèmes informatiques

3.2. Classification des attaques

3.2.1. Attaques agissant au niveau des systèmes logiciels

Les systèmes logiciels constituent un premier niveau d'abstraction à partir duquel un attaquant peut mettre en défaut la sécurité d'un système informatique. Dans ce contexte, le terme système logiciel doit être pris dans son sens le plus général. Il désigne tous types de logiciels dans un système informatique, couvrant aussi bien les programmes d'application, le système d'exploitation et son noyau, que les logiciels implantés (en anglais, firmware) dans les composants matériels. Une attaque, à ce niveau d'abstraction, repose alors soit sur l'utilisation d'une fonctionnalité logicielle légitime du système (éventuellement accessible grâce à une erreur dans la configuration logicielle), soit sur l'exploitation d'une fonctionnalité logicielle vulnérable à des fins malveillantes. La présente sous-section discute de ces deux vecteurs d'attaque.

3.2.2. Attaques agissant au niveau des systèmes matériels

Les systèmes matériels constituent un second niveau d'abstraction à partir duquel il est possible de nuire à la sécurité d'un système informatique. Bien qu'il soit plus simple pour un attaquant de cibler directement les systèmes logiciels, nous observons qu'actuellement de plus en plus d'attaques s'appuient sur les systèmes matériels pour impacter indirectement le système logiciel. Nous discernons deux raisons principales à cela. La première est que le fonctionnement des systèmes logiciels repose sur les systèmes matériels. Ainsi, corrompre un système matériel signifie potentiellement corrompre tous les systèmes logiciels qui en dépendent pour leur exécution. Le fait

que les systèmes matériels sont souvent soumis à des restrictions moins drastiques que les systèmes logiciels constituent une seconde raison. En effet, au niveau du matériel, il n'existe plus de notion de privilèges, d'isolation de processus, etc. Une attaque au niveau des systèmes matériels agit alors soit sur les fonctionnalités matérielles qui ont été implémentées en logique câblée, soit sur leur configuration matérielle. Une telle attaque peut être mise en œuvre de différentes façons, présentées dans les trois sous-sections suivantes.

3.2.3. Attaques agissant au niveau des canaux de communication

Les canaux de communication constituent un autre niveau d'abstraction à partir duquel un attaquant peut mettre en défaut la sécurité d'un système informatique. La notion de canal de communication désigne ici tout type de médium de transmission d'information dont le rôle est d'acheminer des messages entre des systèmes qui interagissent, et couvre aussi bien les canaux de communication physiques que les canaux de communication logiques (par exemple, les mémoires partagées, les fichiers partagés, etc.). Une attaque, à ce niveau d'abstraction, repose alors sur des vecteurs d'attaque liés aux messages échangés entre les systèmes logiciels ou matériels qui interagissent. Nous distinguons alors quatre vecteurs d'attaque possibles qui nécessitent pour l'attaquant un accès au canal de transmission : la destruction, la modification, la captation et l'insertion de messages, ces messages pouvant être soit cohérents, soit incohérents par rapport au protocole de communication. La suite de cette sous-section présente de façon succincte ces vecteurs d'attaque.

3.2.4. Attaques agissant au niveau des canaux auxiliaires

L'environnement d'un système informatique constitue un dernier niveau d'abstraction à partir duquel un attaquant peut porter atteinte à la sécurité de ce système. En particulier, un attaquant peut agir au niveau des canaux auxiliaires, c'est-à-dire les canaux autres que ceux généralement utilisés pour la transmission d'information. Une attaque, à ce niveau d'abstraction peut alors consister à analyser les canaux de fuite ou à injecter des fautes dans le système informatique via les canaux auxiliaires. La suite de cette sous-section explore ces vecteurs d'attaque.

Toutefois, notre intention n'est pas de faire un examen exhaustif des techniques existantes, mais plutôt de mettre en évidence la philosophie qui sous-tend les principales attaques. Aussi, seuls les canaux auxiliaires les plus usités seront évoqués ici.

3.3. Méthodes et techniques pour l'élimination des fautes

Comme précédemment indiqué, la conception et la réalisation d'un système informatique sûr de fonctionnement passent par l'utilisation combinée d'un ensemble de méthodes qui sont classées en quatre moyens : la prévention, l'élimination, la tolérance et la prévision des fautes. La présente section s'intéresse plus particulièrement à l'élimination des fautes, qui vise à réduire le nombre et la sévérité des fautes. Elle est réalisée pendant la phase de développement d'un système par vérification, diagnostic et correction, ou pendant sa phase opérationnelle par maintenance. **[23]**

La vérification consiste à déterminer si le système satisfait des propriétés, appelées conditions de vérification **[29]** ; si ce n'est pas le cas, deux autres étapes doivent être entreprises : diagnostiquer la ou les fautes qui ont empêché les conditions de vérification d'être remplies, puis apporter les corrections nécessaires. Après correction, le processus doit être recommencé afin de s'assurer que l'élimination des fautes n'a pas eu de conséquences indésirables.

Les fautes présentes dans un système peuvent être révélées par différentes techniques qui sont classées selon qu'elles impliquent ou non l'activation du système (**figure 1.16**). La vérification d'un système sans activation réelle est dite statique, par opposition à la vérification dynamique qui nécessite son activation.

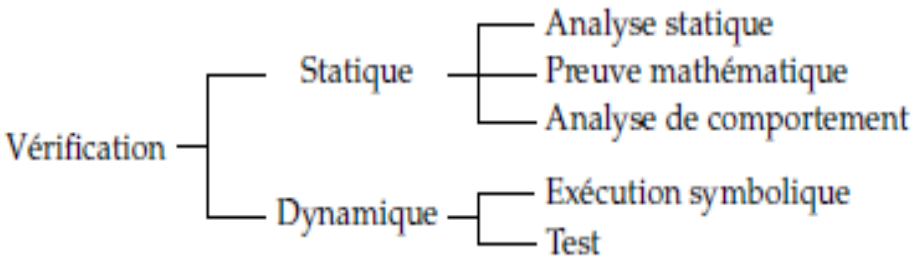

Figure 4.4 – Classification des techniques de vérification

3.3.1. Analyse statique

L'analyse statique peut être manuelle, ou automatique. L'analyse statique manuelle correspond aux techniques de revue ou d'inspection prévues tout au long du cycle de développement du système. **[30][31][32][33][34]**

Une équipe de personnes se réunit pour analyser en détail les documents (par exemple, spécification, ou dossier de conception générale ou détaillée, ou code-source) relatifs à la phase en cours. L'examen de ces documents est guidé par des listes de contrôle (en anglais, check-list) (contenant des questions à se poser, des standards à respecter, etc.) et donne lieu à des questions ou remarques qui sont discutées avec les auteurs du document. L'analyse statique automatique regroupe tous les contrôles effectués à l'aide d'outils informatiques – des compilateurs aux analyseurs les plus évolués. Ces outils fournissent des caractéristiques du code (mesures de complexité, références croisées, chemins d'exécution, etc.) qui facilitent le travail d'analyse statique manuelle et signalent certaines anomalies structurelles (variables non-initialisées, interfaces incompatibles entre composants, utilisations incohérentes de variables globales, code mort, etc.). À cet effet, les algorithmes mis en œuvre dans ces outils peuvent se baser sur des heuristiques simples ou sur des modèles ou théories plus complexes tels que les graphes de contrôle ou l'interprétation abstraite. **[35][36]**

3.3.2 Preuve mathématique

Une preuve mathématique est une suite finie d'inférences dans un système formel, c'est-à-dire dans un formalisme ayant une syntaxe et une sémantique s'appuyant sur des fondements mathématiques associés à des règles de manipulation permettant d'effectuer des transformations et des vérifications. La preuve du programme dont est issu un système logiciel ou matériel par rapport à sa spécification nécessite une formalisation du problème de vérification : la description du système, les hypothèses formulées sur son environnement et les propriétés attendues doivent être exprimées dans un langage formel qui s'appuie sur un cadre logique bien défini.

La plupart des méthodes qui concernent les programmes séquentiels ont recours à une approche basée sur la sémantique axiomatique. Cette sémantique, dont les fondements remontent à **[37][38][39]**, consiste à définir une logique mathématique permettant de prouver des propriétés sur des programmes écrits dans un langage de programmation donné.

3.3.3 Analyse de comportement

L'analyse de comportement est basée sur des modèles comportementaux du système, déduits de la spécification, ou des dossiers de conception, ou des codes-sources **[40][41][42]**.Cette analyse met en jeu des simulations de son comportement, modélisées à partir de graphes, d'automates à états finis, de tables de décision, de réseaux de Pétri, etc. Il est possible de vérifier des propriétés de cohérence, complétude, vivacité, temps de réponse, etc. dès lors qu'une sémantique formelle est associée au modèle utilisé. La vérification de modèles(en anglais, model checking), proposée par **[43] [44]**, constitue un cas particulier des techniques d'analyse de comportement. Il s'agit d'un ensemble de techniques de vérification automatique de propriétés temporelles sur des systèmes réactifs (par exemple, protocoles de communication, composants électroniques, etc.), c'est-à-dire un système qui est en interaction permanente avec son environnement et dont le rythme est imposé par cet environnement. Elle nécessite de formuler les spécifications dans une logique temporelle propositionnelle et de modéliser le système sous la forme d'un graphe orienté, constitué d'états et de transitions. L'outil de model checking (ou model checker) parcourt ensuite de façon exhaustive le graphe en s'assurant qu'une propriété est vérifiée pour chacun de ses états et retourne un contre-exemple si la propriété a été violée dans au moins un de ses états.

3.3.4 Exécution symbolique

L'exécution symbolique de programme consiste à exécuter un programme en lui soumettant des valeurs symboliques en entrée : par exemple, si une entrée X est un nombre entier positif, nous pouvons lui affecter une valeur symbolique a qui représente l'ensemble des valeurs entières supérieures à 100. Une exécution symbolique consiste alors à propager les symboles sous forme de formules au fur et à mesure de l'exécution des instructions. Elle fournit comme résultats les expressions symboliques obtenues pour les sorties. En pratique, les limites auxquelles se heurte cette technique sont nombreuses : dans l'exemple précédent, comment connaitre le résultat d'une condition de branchement portant sur une valeur précise de X (par exemple, X = 200), résultat qui conditionne la suite des instructions à exécuter ? Que représente l'élément d'indice X dans un

tableau lorsque X a la valeur a ? Comment maîtriser l'explosion de la taille et de la complexité des formules obtenues ?

3.3.5 Test

Le test est certainement la technique de vérification la plus largement répandue **[45]**.Il consiste à exécuter un programme en lui fournissant des entrées values, les entrées de test, et à vérifier la conformité des sorties par rapport au comportement attendu. Sa mise en œuvre nécessite de résoudre deux problèmes : le problème de la sélection d'entrées de test, et le problème de l'oracle **[46]**, autrement dit comment décider de l'exactitude des résultats observés. Sauf cas trivial, le test exhaustif est généralement impossible et on est amené à sélectionner de manière pertinente un (petit) sous-ensemble du domaine d'entrée. Cette sélection peut s'effectuer à l'aide de critères de test liés à un modèle de la structure du système. Nous parlons alors de test structurel. Elle peut également s'effectuer à l'aide de critères de test liés à un modèle des fonctions que doit réaliser le système. On parle alors plutôt de test fonctionnel.

Quel que soit le critère de sélection, la génération des entrées peut être déterministe ou probabiliste.

Dans le premier cas qui définit le test déterministe, les entrées sont déterminées par un choix sélectif de manière à satisfaire le critère de test retenu. **Dans le deuxième cas** qui définit le test statistique ou aléatoire, les entrées sont générées de manière aléatoire selon une distribution probabiliste sur le domaine d'entrée, la distribution et le nombre des données de test étant déterminé sà partir du critère retenu **[47]**. Ces deux types de génération des entrées de test, déterministe et aléatoire, sont complémentaires **[48][49]**,et devraient être utilisés à toutes les étapes de test. Il est important de rappeler le rôle déterminant du choix de la distribution des probabilités sur le domaine d'entrée, dont dépend fortement l'efficacité du test statistique : la distribution doit être choisie en fonction du critère de test pour permettre une bonne couverture, structurelle ou fonctionnelle, du système.

En ce qui concerne le problème de l'oracle, un dépouillement automatisé des résultats de test est toujours souhaitable, et devient indispensable lorsqu'un grand nombre d'entrées ont été sélectionnées. Les solutions les plus satisfaisantes sont basées sur l'existence d'une spécification formelle du programme sous test. La spécification est alors utilisée pour déterminer les résultats attendus, soit lors de la sélection des entrées de test, soit a posteriori. D'autres solutions d'oracle partiel peuvent être déterminées au cas par cas, en réalisant des contrôles de vraisemblance sur les résultats de test (contrôle de cohérence entre différentes données, contrôle d'appartenance à une plage de valeurs, etc.).

4. Des réseaux de capteurs vulnérables

Dans l'IoT, on considère souvent des objets (capteurs/actuateurs) qui ont des contraintes matérielles et logicielles qui ne leur permettent pas de se connecter directement au réseau Internet. Ils s'y connectent à travers une gateway (passerelle). En effet, d'un côté, l'Internet n'est pas dimensionné pour gérer l'adressage de milliards d'objets connectés. D'un autre côté, TCP IP/V6 est un protocole trop lourd pour être exploité par les capteurs. Aujourd'hui, TCP IP V6 est utilisé via le protocole 6LowPAN (IPv6 Low power Wireless Personal Area Networks) qui est exploité au-dessus des protocoles réseaux 802.15.4.

Les gateways jouent le rôle d'intermédiaire pour connecter l'objet à internet et envoyer ses données au cloud. Un exemple de ces gateways sont les routeurs domestiques, les téléphones mobiles, le raspberryPi, l'Arduino. Ces gateways fournissent ce qui est nécessaire en termes de connectivité, de sécurité et de management des appareils. Les gateways traduisent aussi les protocoles propriétaires (exemple zigbee, BLE) au réseau Internet et certaines peuvent jouer le rôle d'agrégateurs de réseaux.

4.1. Spécificités des réseaux de capteurs sans fil

Les réseaux de capteurs sans fil sont des réseaux ad-hoc spécifiques **[50]** avec un nombre de nœuds plus conséquents, une énergie limitée et une puissance de calcul plus faible que les réseaux ad-hoc classiques. Ce sont ces particularités que nous introduisons dans la partie suivante.

4.1.1 Topologie

La topologie que l'on retrouve classiquement au sein des réseaux de capteurs sans fil est un ensemble de nœuds (chaque nœud représentant un capteur) qui sont déposés de manière hétérogène sur une zone ou des objets voir des individus mouvants. Tous ces nœuds communiquent entre eux, chaque nœud peut communiquer avec les autres nœuds qui sont situés dans sa zone de couverture.

Les réseaux de capteurs sans fil sont le plus souvent reliés à une ou plusieurs bases. Ces bases ont pour mission de récupérer les informations circulant sur le réseau, et de les stocker ou bien de les envoyer directement via une liaison internet ou une liaison GSM.

Ces bases peuvent être par exemple un ordinateur portable ou un capteur de puissance plus importante que les autres nœuds classiques. Elles peuvent avoir un rôle de contrôleur du réseau et elles font souvent le lien entre l'utilisateur et le réseau.

4.1.2. Routage

Pour limiter le nombre de communications coûteuses en énergie, les réseaux de capteurs sans fil utilisent des protocoles de routage efficaces **[51]**. Une solution souvent utilisée est la clustérisassions, qui divise le réseau en plusieurs clusters. Dans chacun de ces clusters, un nœud

maître (cluster-head) est élu et aura pour mission de récupérer les informations des nœuds du cluster dont il a la charge pour les transmettre aux autres clusters et inversement.

Le choix du nœud maître sera fait en désignant par exemple le nœud avec l'énergie la plus importante, pour augmenter la vie du réseau.

D'autres problèmes de routage doivent aussi être pris en compte pour limiter le nombre de communications comme les problèmes d'implosion ou de chevauchement qui sont expliqués dans. **[52]**

4.1.3. La tolérance aux fautes

Dans les réseaux de capteurs sans fil, un ou plusieurs capteurs peuvent ne pas fonctionner correctement. En effet les capteurs sont des entités sensibles aux altérations d'états comme des phénomènes climatiques (humidité, chaleur, électromagnétisme) ou du fait d'une batterie faible.

Dans ce cas de figure, le réseau doit être capable de détecter ce type d'erreur est d'y remédier, en cherchant par exemple à modifier ses tables de routage pour trouver un autre chemin permettant de transmettre l'information et de maintenir le réseau toujours opérationnel.

De la même manière, les capteurs doivent pouvoir détecter des capteurs défaillants qui envoient des informations erronées du fait de leur état.

4.1.4. Mise à l'échelle

Le nombre de capteurs utilisés dans les réseaux de capteurs sans fil peut varier de quelques entités à plusieurs dizaines de milliers. C'est d'ailleurs la principale utilité des réseaux de capteurs qui doivent pouvoir s'auto organiser à une grande échelle et être efficace quel que soit le nombre. Pour cela les protocoles des réseaux de capteurs sans fil doivent être capables de fonctionner et de s'adapter selon le nombre de nœuds.

4.1.5. Une énergie limitée

Les capteurs sont équipés de batteries avec une énergie limitée (plusieurs jours à quelques années). De plus, les réseaux de capteurs sans fil quand ils sont déployés, le sont souvent dans des zones difficiles d'accès pour l'homme. Il est donc difficile de pouvoir changer les batteries des capteurs. Si le nombre des capteurs dépasse la centaine d'entités, il est encore plus difficile d'intervenir pour trouver le capteur défaillant et changer sa batterie.

Les capteurs sont en général déployés pour ne plus être modifiés.

La consommation de l'énergie des réseaux de capteurs sans fil doit être la plus préservée possible. Dans ce but, les capteurs actuels ont des périodes de veille durant leur temps d'inactivité pour préserver leur batterie.

Les communications sont les actions les plus coûteuses en termes d'énergie. Les calculs le sont, mais dans une moindre importance. Il est donc fortement nécessaire de limiter le nombre de communications entre capteurs et si possible le nombre de calculs.

4.1.6. Faible puissance de calcul

Malgré les progrès récents dans la fabrication de capteurs de plus en plus puissants, les capteurs actuels souffrent d'un manque de puissance de calcul (par exemple seulement16 Mhz de puissance et 128Koctets de mémoire programmable pour un capteur MicaZ).

Cette faible puissance ne permet pas d'utiliser des algorithmes complexes dans les réseaux de capteurs sans fil, et particulièrement dans la cryptographie poussée.

De plus la vocation des capteurs sans fil est d'être en très grand nombre et leur utilisation dans des applications avec un nombre de nœuds élevé nécessite l'utilisation de capteurs bons marchés, ce qui impliquent des capteurs avec une puissance de calcul très faible.

La faiblesse de la puissance de calcul est aussi préjudiciable pour le temps de réponse du réseau. Si l'on demande à un capteur d'effectuer de nombreux calculs, sa réactivité va sensiblement se détériorer.

4.2. Présentation des attaques

Les différentes spécificités des réseaux de capteurs sans fil (énergie limité, faible puissance de calcul, utilisation des ondes radio, etc..) les exposent à de nombreuses menaces.

Si certaines de ces menaces peuvent se retrouver dans les réseaux ad-hoc d'autres sont spécifiques aux réseaux de capteurs sans fil et s'attaquent plus particulièrement à l'énergie limitée des capteurs.

On parlera d'attaque active si un attaquant modifie l'état du réseau, et d'attaque passive dans le cas où il ne cherchera qu'à l'écouter.

4.2.1. Destruction ou vol

Les plus élémentaires des attaques actives dans les réseaux de capteurs sans fil sont le volou la destruction des capteurs. Les capteurs sont déployés dans des zones qui ne peuvent être toujours surveillées. Ainsi une personne physique seule peut subtiliser un ou plusieurs capteurs, voire peut les détruire. Si un capteur est détruit, le réseau doit être capable de s'adapter à la nouvelle situation et éviter d'être divisé en plusieurs sous-réseaux incapables de communiquer entre eux.

De plus, un nœud volé, peut divulguer certaines informations à un attaquant. Il peut tout aussi bien être reprogrammé et être réinséré dans le réseau et ainsi devenir un nœud malicieux, fonctionnant en tant que nœud espion comme expliqué dans **[53]**, **[54]** et **[55]**.

4.2.2. Attaque spécifique au type de capteur

Ce type d'attaque dépend du type de capteur utilisé sur le réseau.

Un attaquant va modifier de manière physique le comportement du capteur. Il peut par exemple allumer une flamme devant un capteur thermique ou bien allumer une lampe devant un capteur de luminosité. Le but est de tromper le capteur, et ainsi d'envoyer ou d'enregistrer de fausses informations sur le réseau, ou bien tout simplement de faire réagir assez longtemps un nœud ou le réseau pour qu'ils consomment leur énergie.

4.2.3. L'écoute passive

Cette attaque consiste à écouter le réseau et à intercepter les informations circulant sur le médium. Cette attaque est facilement réalisable si les messages circulant sur le réseau sont en clair. Par ailleurs cette attaque est difficile à détecter, car comme elle est passive, elle ne modifie pas l'activité du réseau.

4.2.4. Brouillage radio

Un attaquant va envoyer des ondes sur la même fréquence que le réseau de capteurs sans fil **[56]**. Ainsi les nœuds ne pourront plus communiquer car le médium est saturé par la brouillage radio.

4.2.5 L'injection de messages

L'attaquant va chercher par divers moyens à injecter des messages dans le réseau. Le but peut être de faire circuler de fausses informations ou tout simplement de saturer le réseau.

4.2.6. Flooding

Un attaquant va utiliser un ou plusieurs nœuds malicieux ou un dispositif particulier avec une puissance d'émission forte, pour envoyer régulièrement des messages sur le réseau pour le saturer.

On est en présence d'une attaque active qui est de même type que les attaques de type déni de service dans les réseaux classiques **[39]**.

4.2.7. Hello Flooding

Les protocoles de découvertes sur les réseaux ad-hoc utilisent ce qu'on appelle des messages de type HELLO pour s'insérer dans un réseau et pour découvrir ses nœuds voisins.

Dans une attaque dite de HELLO Flooding, un attaquant va utiliser ce mécanisme pour saturer le réseau et consommer son énergie.

Dans **[57]**, on trouve un exemple, représenté par la figure 1, d'un nœud malicieux X avec une connexion puissante qui lui permet d'envoyer à un grand nombre de nœuds des messages de type

HELLO, de manière continue. Les nœuds voisins V vont alors essayer de lui répondre, même s'ils sont situés à des distances qui ne permettent pas d'atteindre le nœud malicieux. A force de tenter de répondre à ces messages ils vont petit à petit consommer l'intégralité de leur énergie.

4.2.8. La privation de mise en veille

Cette attaque active a pour but de priver un capteur de se mettre en veille par différents moyens **[58]**. Le capteur s'il ne peut plus se mettre en veille va consommer très rapidement sa batterie, jusqu'à se retrouver hors service.

4.2.9. Insertions de boucles infinies

Un attaquant va modifier le routage du réseau avec un ou plusieurs noeuds malicieux,dans le but d'envoyer des messages qui vont être routés en boucles infinies et vont doncconsommer l'énergie du réseau.

4.2.10 L'altération de message

Un nœud malicieux va récupérer un message et l'altérer, en lui ajoutant des fausses informations (sur le destinataire, l'émetteur ou les données), en le modifiant ou bien en détruisant des paquets pour rendre incompréhensible le message.

4.2.11. Ralentissement

Un attaquant peut programmer des nœuds malicieux qui seront comme des agents dormants qui n'auront que pour but de ralentir l'information (par exemple avec une attaque de type trou gris).

4.2.12 Attaque du trou noir (black holeattack)

L'attaque du trou noir consiste tout d'abord à insérer un nœud malicieux dans le réseau **[59]**.

Ce nœud, par divers moyens, va modifier les tables de routage pour obliger le maximum de nœuds voisins à faire passer l'information par lui. Ensuite comme un trou noir dans l'espace, toutes les informations qui vont passer en son sein ne seront jamais retransmises.

4.2.13. Sybil attack

Une attaque de type "Sybil attack" **[60]** consiste à ce qu'un capteur malicieux se fasse passer pour plusieurs capteurs. Il va ainsi pourvoir modifier la table de routage qui deviendra caduque. Un nœud malicieux qui peut se faire passer pour plusieurs nœuds peut gagner un avantage important pour une élection de nœud maître par exemple.

4.3. Mécanismes de sécurité

Pour contrer les attaques qui menacent les réseaux de capteurs sans fil, plusieurs équipes de recherche tentent de trouver des solutions appropriées. Ces solutions doivent bien sûr prendre en compte les spécificités des réseaux de capteurs sans fil. Il faut donc trouver des solutions simples qui permettent de sécuriser le réseau tout en consommant le moins d'énergie possible et adapter ces solutions à une puissance de calcul faible.

Dans l'éventail de ces solutions, on trouve des mécanismes tels que le partitionnement de données, l'utilisation de méthodes cryptographiques adaptées, la détection d'intrus par localisation ou bien encore l'indice de confiance.

4.3.1. Le partitionnement des données

[61] et **[62]** offrent une solution pour empêcher la récupération d'information dans les réseaux de capteurs sans fil par le partitionnement des données. Comme son nom l'indique le but est de découper l'information en plusieurs parties.

Si un capteur cherche à envoyer une information, celui-ci va la découper en plusieurs paquets de taille fixe. Chaque paquet sera ensuite envoyé sur des chemins différents, c'est à dire qu'elles ne passeront pas par la même route et donc les mêmes nœuds. Ces paquets seront finalement reçus par la base, qui pourra ensuite les rassembler pour pouvoir reproduire l'information. Ce mécanisme oblige un attaquant à récupérer l'ensemble des paquets s'il veut pouvoir lire l'information. Il doit aussi être capable d'écouter l'ensemble du réseau, pour récupérer les différents paquets qui circulent sur des chemins différents.

Un exemple de cette solution est représenté par la figure 5, où un capteur A divise un message en 3 paquets qui vont suivre respectivement 3 chemins différents.

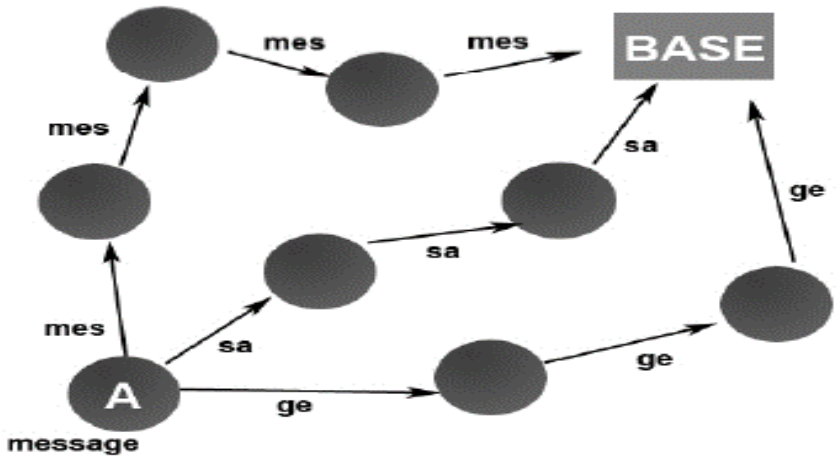

Figure 4.5 : Exemple de partitionnement

Cette solution oblige un agresseur à écouter l'ensemble du réseau et à récupérer l'ensemble des messages pour parvenir à récupérer l'information. Cependant cette solution augmente considérablement la consommation d'énergie (avec un risque de surcharge de traitement),car elle sollicite un nombre de nœuds plus importants.

4.3.2. La cryptographie

Comme nous l'avons expliqué auparavant, il n'est pas possible dans les réseaux de capteurs sans fil d'utiliser des méthodes de cryptographie complexes. La faible puissance des capteurs ne le permet pas, et quand elle le permet, le temps de calcul est trop long.

Cependant, il est possible d'utiliser des techniques de cryptographie simple avec des clés symétriques comme montré dans **[63]**.

Quatre types de cryptographie sont ainsi utilisés:

a. Clé globale:

Une clé est partagée par l'ensemble du réseau. Pour envoyer un message, l'information est chiffrée avec cette clé. Une fois le message reçu, le message peut être déchiffré avec cette même clé (principe de la clé symétrique). C'est la solution la moins coûteuse en termes d'énergie, mais avec la sécurité la moins importante. Si un agresseur récupère la clé, il peut déchiffrer tout le réseau.

b. Clé partagée par paire de nœuds :

Chaque nœud possède une clé différente pour communiquer avec un nœud voisin qui partage cette clé. Ainsi si un nœud possède" voisins, il aura "n" clés à stocker pour pouvoir communiquer avec ses voisins. Dans cette solution, un nœud qui cherche à envoyer un message, doit l'encrypter avec la clé du voisin qui recevra l'information. Le nœud voisin devra déchiffrer l'information pour la chiffrer à nouveau avec la clé qui correspond au destinataire suivant. C'est la solution cryptographique la plus sécurisée (l'agresseur doit récupérer chaque clé par paire de nœuds pour avoir accès à toute l'information), mais aussi la plus coûteuse en terme d'énergie et de latence. Chaque nœud intermédiaire doit déchiffrer le message du prédécesseur, puis le chiffrer avant de l'envoyer au nœud suivant.

c. Clé partagée par groupe de nœuds:

Dans ce cas de figure, chaque groupe ou cluster partage une clé en commun qui lui permet de communiquer à l'intérieur du groupe.

Les nœuds maîtres communiquent entre eux avec, soit une clé commune à tous les clusters heads, soit une clé commune par paire de cluster head. Cette solution est une solution hybride des deux premières techniques de chiffrement et apporte un compromis entre sécurité et consommation d'énergie.

d. Clé individuelle:

Dans cette solution chaque nœud possède une clé personnelle pour chiffrer son information. Cette clé n'est connue que de la base. Ainsi un message envoyé par ce nœud circulera de manière cachée sur le réseau jusqu'à atteindre la base. Si cette solution est intéressante en termes de sécurité, elle n'apporte qu'une possibilité de communication sécurisé entre un nœud et la base, mais pas entre nœuds.

4.3.3. Génération

Une solution proposée par **[64]** consiste à utiliser une clé de génération. A chaque période ou génération, la base envoie une nouvelle clé à l'ensemble du réseau.

Cette clé sert de certificat à chacun des nœuds, pour prouver son appartenance au réseau. Si un nœud non identifié tente de rentrer dans le réseau de capteurs sans fil et qu'il ne possède pas cette clé de génération, il ne pourra être accepté en son sein.

Un autre intérêt de cette technique est qu'elle permet de limiter les attaques de substitution d'un capteur et de sa reprogrammation pour être réinjecté dans le réseau.

Si ce nœud est subtilisé à l'instant 0 avec la clé de génération K(0), le temps qu'un attaquant le reprogramme pour le remettre dans le réseau il se sera écoulé un temps "x".

Quand le capteur sera repositionné dans le réseau, la nouvelle clé de génération sera alors K(x). Le nœud malicieux demandera à ses nœuds voisins de rentrer dans le réseau avec la clé K(0) et non pas K(x), car il n'a pas pu recevoir la nouvelle clé. Comme K(0)!=K(x), les nœuds voisins n'accepterons pas sa requête et le nœud malicieux ne pourra pas rentrer dans le réseau.

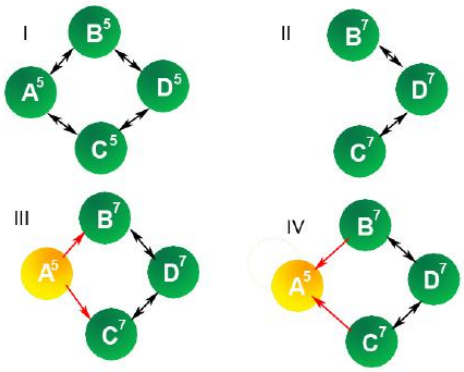

Fig. 4.6 : Détection de nœud malicieux par clé de génération

Un exemple est donné par la figure 6, où quatre capteurs A, B, C, D font partie d'un réseau de capteurs qui communiquent par clés symétriques par paire de nœuds. A l'étape I, les capteurs ont pour clé de génération 5. A l'étape II, le nœud A est subtilisé par un attaquant, et pendant son absence sur le réseau, la base transmet une nouvelle clé de génération 7. A l'étape III, le capteur A reprogrammé et réinséré dans le réseau fait une demande d'insertion dans le réseau à B et C. A

l'étape IV, les nœuds B et C refusent la demande de A, car en comparant leur clé de génération, ils se sont aperçus qu'elles ne correspondaient plus.

Cette technique est peu coûteuse en terme d'énergie et facile à déployer. Cependant elle ne s'adresse qu'à des réseaux fermés, qui ne peuvent pas accepter de nouveaux nœuds.

4.3.4. Localisation

Un mécanisme utilisé pour détecter les nœuds malicieux et particulièrement des attaques de type trou de ver, consiste à utiliser une technique de localisation géographique, comme proposé par **[65]** et **[66]**.

Pour cette solution, le réseau de capteurs sans fil doit être équipé de capteurs balises (beaconnode), qui sont des capteurs qui connaissent leur position géographique, par exemple au moyen d'un équipement GPS.

Avec la localisation, si un capteur demande à entrer dans le réseau, les capteurs balises qui vont recevoir cette demande vont pouvoir estimer sa localisation par rapport à son domaine d'écoute. Les capteurs balises vont ensuite quadriller leur zone d'écoute respective, et chaque nœud qui a reçu la demande d'insertion dans le réseau va voter pour une zone du quadrillage qu'il est capable d'entendre. La zone qui obtiendra le plus grand nombre de voix sera considérée comme la zone où est censé se trouver le nouveau capteur.

La figure 7 montre un exemple de vote entre 4 capteurs balises A, B, C et D qui ont quadrillé leur zone d'écoute respective et qui ont chacun voté pour chaque zone de quadrillage. A la fin du vote, ils peuvent estimer la position du capteur recherché. Ce dernier doit potentiellement se trouver dans la zone avec le maximum de votes, c'est à dire dans l'exemple, la zone avec 3 votes.

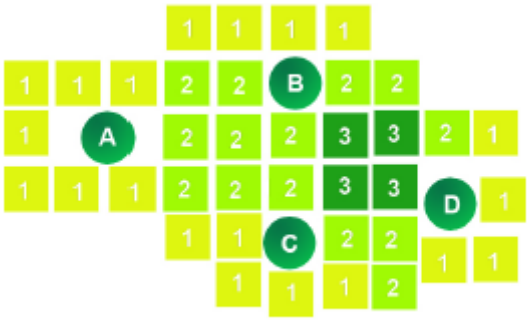

Fig. 4.7: Exemple de localisation avec des capteurs de type beacon

4.3.5 L'indice de confiance et la réputation

Une solution proposée par **[67]**, [**68**], **[69]**, **[70]**, **[71]** et **[72]** consiste à utiliser les mécanismes de confiance et de réputation que l'on peut trouver dans les réseaux pair à pair **[73]**, les réseaux de communauté ou bien encore dans les sites marchands comme EBay.

Dans ce type de réseau tout comme dans les réseaux de capteurs sans fil, il est difficile de savoir, au vu du nombre de nœuds, quel nœud peut être un nœud malicieux. Pour le détecter et conserver l'intégrité du réseau, chaque nœud du réseau va surveiller ses nœuds voisins et leurs actions au cours du temps. En fonction des actions réalisées par ses nœuds voisins, un nœud va augmenter une note de l'indice de confiance de ces nœuds, basée sur sa réputation. Si un nœud ne répond jamais à une requête, son indice de confiance va diminuer, de la même manière que si ce nœud retransmet toujours correctement l'information qu'on lui a demandé de transmettre, son indice de confiance va augmenter.

A l'aide de ces indices de confiance, un nœud va alors choisir le routage le plus adapté pour transmettre son information. Contrairement à des protocoles classiques de le nœud chercherait le chemin le plus rapide en nombre de sauts ou de distance géographique, il va choisir ici de transmettre son information via les nœuds avec les indices de confiance les plus élevés, en d'autre terme, la route qui lui semble la plus sûre.

Ces techniques permettent d'éliminer du routage traditionnel les nœuds qui sont potentiellement dangereux, et empêcher ainsi l'information de passer par ces nœuds.

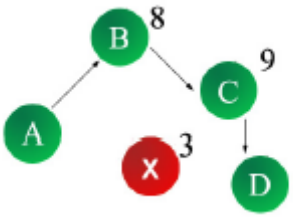

Fig. 4.8 : Choix de routage par réputation

Ce mécanisme est représenté par la figure 8, où un nœud A doit transmettre une information à un nœud D. Au lieu de passer par le chemin le plus court qui passe par X, qui est un nœud avec un indice de confiance faible de 3 (sur une note de 10), et donc est potentiellement un nœud à risque, le nœud A va transmettre l'information par les nœuds B et C qui avec des indices de confiance de 8 et 9 proposent le chemin le plus sûr.

Cette solution peut être jumelée avec un mécanisme de surveillance entre voisins proches nommé mécanisme du chien de garde (watchdog) **[74]**, où pour chaque communication entre deux nœuds A et B, un nœud intermédiaire C, situé dans la zone de communication, est chargé de surveiller que cette communication a bien été effectuée, comme représentée dans la figure 9.

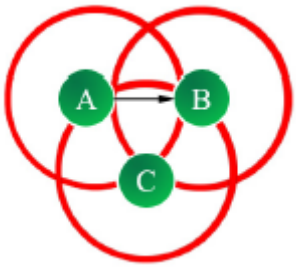

Fig. 4.9 : Exemple de chien de garde

Les solutions basées sur l'indice de confiance sont peu coûteuses en termes d'énergie et permettent, selon le type de sécurité voulu, de ne pas avoir recours à la cryptographie.

Cependant pour des réseaux qui demandent une sécurité maximale, elles ne sont pas toujours adaptées. Ainsi un nœud malicieux qui enregistrerait des informations sur le réseau et, par ailleurs, se comporterait de manière normale, est difficilement détectable.

5. Une plate-forme sécurisée pour éviter les attaques

Réduire la surface d'exposition des objets connectés aux attaques est une tâche complexe. Elle requiert une connaissance architecturale de la chaine de valeur qui relie les objets au cloud. Il faut s'intéresser aux objets eux-mêmes, à leurs capteurs et processeurs, aux réseaux locaux et distants, aux protocoles de tout niveau, puis aux serveurs, à leurs logiciels et aux traitements des données qui y sont réalisés. Les besoins sont bien connus depuis des années (cf "Security needs in embeddedsystems" paru… en 2008). De nombreuses sociétés proposent des outils permettant de sécuriser telle ou telle partie de la chaine de valeur mais elles se positionnent depuis assez peu de temps dans les objets connectés.

Avec les technologies existantes, la sécurité est une affaire de méthodes et de processus. C'est ce qu'expliquait très bien Yann Allain de la société de conseil **Opale Security** à la conférence de Toulouse. Il expliquait le besoin de bien sécuriser toute la chaine de valeur et pas simplement l'objet lui-même. Il évoquait la menace globale provenant de l'absence de chiffrage des données de nombreux objets connectés, du manque de web sécurisé, du manque de confidentialité des données, de la sécurité des mises à jour des firmware des objets, des couches électroniques. Il mettait en évidence le fait que nombre d'industriels ne savent pas évaluer la robustesse de leurs produits et que les outils de tests ne sont pas encore répandus dans l'IOT. Il évoquait des standards de tests de sécurisation d'objets, aussi bien le **JTAG** que **l'OWASP IoT.** [**Web35**]

5.1. Sécurité hard et soft pour IOT

La prise de conscience de l'importance de la dimension sécuritaire s'accompagne souvent du constat d'un manque de maîtrise, ce qui freine nombre de projets. Les questions de sécurité souffrent en effet d'être partagées entre différents acteurs au sein de l'entreprise et d'un déficit

d'interlocuteurs réunissant l'ensemble des compétences nécessaires et capables d'apporter une réponse globale et intégrée.

La sécurité de l'IoT repose en effet sur six piliers :

- La sécurisation des capteurs et de leur fonctionnement
- La confidentialité et l'intégrité des données en transit
- La sécurisation des données stockées
- La sécurisation des accès à l'information
- Une approche « chip to cloud » de la sécurité
- Une stratégie de normalisation appropriée.

La stratégie de sécurité doit toutefois tenir compte des spécificités de l'IoT, que ce soient les protocoles à faible consommation longue portée (LoRa, Sigfox…), adaptés aux systèmes de capteurs disséminés sur des objets dépourvus d'alimentation, ou les protocoles à plus faible portée (Wi-Fi, ZigBee, Bluetooth Low Energy…), qui peuvent s'intégrer à des appareils électriques et/ou bénéficier du relai d'une passerelle.

5.2. Le transit des données au cœur des enjeux de l'IoT

Aux deux extrémités de la chaîne, les questions de sécurité sont bien connues et appréhendées. Les capteurs, d'une part, et les systèmes de traitement, d'autre part, mettent respectivement en œuvre les dispositifs de sécurité usuels des systèmes embarqués et d'information. En revanche, ce sont tous les enjeux liés au transit des données qui constituent la spécificité de l'IoT. Recueillie, acheminée, traitée, stockée, la donnée passe de mains en mains jusqu'à son exploitation. Il faut donc s'assurer de son intégrité et de sa confidentialité tout au long de son parcours jusqu'à l'utilisateur final. Pour cela, se développent des approches globales de la sécurité des services connecté, dites "chip to cloud", qui visent à maintenir l'intégrité de la chaîne grâce à l'identification des objets qui s'y connectent.

Par ailleurs, pour garantir l'évolutivité et la flexibilité des écosystèmes complexes de l'IoT et de l'IoE, la normalisation est incontournable. Plutôt que d'ajouter de nouvelles normes, la GSM Association, qui regroupe les principaux opérateurs télécoms mondiaux, suggère d'utiliser des standards existants. Les recommandations qu'elle a publiées début 2016 pour sécuriser objets, réseaux et services s'appuient sur des technologies éprouvées en matière de cryptographie, d'API, de signatures logicielles [**Web36]**.

5.3. Traitement et stockage sécurisés dans le Cloud

Plus que jamais, la sécurité reste un critère primordial dans le choix d'un service de stockage en ligne. Aujourd'hui, la plupart des solutions ont intégré le chiffrement des données dans leur technologie. Mais cette intégration n'est pas toujours suffisante.

Pour la sécurité dans le nuage, l'une des erreurs à éviter par les fournisseurs de solutions consiste à consacrer tous leurs investissements à fortifier le centre de traitement des données – bunkers souterrains, gardes armés, contrôles d'accès et dispositifs de sécurité électroniques pour protéger le réseau – en négligeant de sécuriser les points d'accès à distance au centre de données.

Dans le monde de l'Internet des objets, c'est la même problématique, mais avec une multiplication exponentielle des points d'accès dont le nombre peut atteindre… plusieurs milliards. Or les attaquants cherchent toujours un maillon faible dans les systèmes informatiques et les objets non protégés vont devenir une cible de choix pour ces derniers. Car une fois un objet sous leur contrôle, les attaquants s'en servent pour accéder à l'intérieur des centres de données, avec la possibilité de récupération de données très sensibles liées à la santé, aux activités sociales, etc.

A titre d'exemple, lors de la récente cyberattaque menée contre la société Target et sa chaîne de grands magasins implantés sur le territoire américain, les logiciels malveillants (malware) ont été insérés dans les objets du détaillant, c'est-à-dire les terminaux points de vente, et il n'y a pas eu d'infiltration directe au niveau des serveurs de la société ou des serveurs de traitement des paiements.

Par ce biais, les attaquants ont fait main basse sur plus de 100 millions de numéros de cartes de crédit et autres données personnelles. Autre exemple récent : des chercheurs d'une université américaine dans le domaine de la sécurité ont identifié un réseau zombie (botnet) constitué d'une myriade d'appareils électroménagers intelligents (réfrigérateurs, aspirateurs…) dont les micros contrôleurs étaient contrôlés à l'insu de leurs utilisateurs par un groupe de hackers.

Une autre erreur fondamentale commise dans l'informatique ennuage consiste à croire qu'une liaison HTTPS entre le dispositif d'accès et le nuage est suffisante pour protéger les informations traversant le Web. Or, à mesure que l'Internet des objets devient plus complexe, il est impossible aux développeurs d'appréhender le nombre de données qui circuleront sur le Net et de savoir si les divers systèmes sur le parcours sont dignes de confiance. Par exemple, lorsque l'on relie un navigateur à Facebook, l'utilisation du protocole de transmission sécurisée SSL protège certes les informations en transit vers la zone d'accueil (DMZ) de Facebook (hypothèse cependant fragilisée par l'absence d'authentification mutuelle rigoureuse dans la plupart des transactions sur le Web), mais qu'en est-il des données pénétrant dans le nuage de Facebook ? Elles peuvent être transmises à des annonceurs, à des bases de données et à des services Web tierces parties. Il est en fait quasi impossible aux développeurs d'applications embarquées dans les objets de s'assurer de la qualité des contrôles de sécurité mis en place par tous ces acteurs. Conséquence : il devient nécessaire d'adopter une stratégie zéro confiance en estimant que le nuage est intrinsèquement dangereux. Si le système sur lequel on travaille génère des données reliées au nuage, alors il est impératif de protéger ces données, qu'elles s'aient ou non des chances de circuler sur le Web. Exemple :

un équipement de soin médical portable peut crypter les informations générées localement à l'aide d'une clé commandée par son propriétaire et partager celles-ci avec les seuls prestataires besoins de santé qui ont le droit d'y accéder.**[75]**

6. Dimensions de la sécurité de l'IdO

L'IdO est une technologie caractérisée par une forte ubiquité dans le monde physique et une omniprésence autour de ses usagers. Les diverses applications potentielles de l'IdO l'hétérogénéité de ses technologies habilitantes et sa forte dimension humaine et socioéconomique rendent sa sécurité un sujet difficile et complexe. En plus des problèmes de sécurité des technologies qui le constitueront, l'IdO accentue les problèmes de sécurité des personnes qui l'utiliseront, et fait émerger de nouveaux problèmes liés à la sécurité des systèmes sous son contrôle. Comme nous l'illustrons sur la Figure 63, la sécurité et la privacy dans l'IdO peut être abordée de trois angles complémentaires qui reflètent ses dimensions technologique, humaine et systémique.

La protection de la technologie concerne en premier lieu la sécurité des données, des communications et des infrastructures réseaux. Cette protection est nécessaire pour contrarier les attaques classiques et futures sur l'intégrité, l'authenticité et la confidentialité des données, ainsi que les attaques sur les infrastructures réseaux et leurs fonctionnalités.

La protection des personnes concernera la protection de la vie privée des usagers (« privacy ») qui nécessite, en plus des solutions technologiques, une régulation appropriée qui établit les responsabilités en cas de litiges. La protection des systèmes interconnectés et hébergeant les objets de l'IdO, concernera la protection des objets eux-mêmes livrés à ces systèmes et les processus qu'ils contrôleront. **[76]**

7. En conclusion : les piliers essentiels pour sécuriser l'IoT

L'Internet des objets va avoir un impact sur nos interactions avec le monde qui nous entoure. Des milliards d'« objets » communiquent entre eux – des téléviseurs, réfrigérateurs et voitures aux compteurs intelligents, moniteurs de santé et wearables. L'IoT promet un confort sans précédent. Cependant, pour que l'IoT puisse déployer tout son potentiel, il est essentiel de gagner et de conserver la confiance des consommateurs en matière de confidentialité et de sécurité. En effet, les transferts de données liés à l'IoT traceront le portrait de chacun d'entre nous. L'enjeu est de sécuriser ces informations. **[web34]**

De nombreux moyens sont à la disposition d'un hackeur pour accéder aux fonctionnalités ou aux données d'un appareil connecté. Les trois principales cibles de « piratage » sont les suivantes : l'appareil, l'infrastructure cloud et le réseau.

A notre avis, il existe trois piliers essentiels pour sécuriser un appareil IoT et garantir la sécurité des données stockées et des données mobiles.

7.1. Pilier numéro un - La sécurisation de l'appareil:

Les milliards d'appareils connectés vont augmenter l'utilisation des logiciels et des données dans les actifs des entreprises et les gadgets grand public. Cela offre de nouvelles vulnérabilités aux pirates malveillants. Les solutions logicielles embarquées de Gemalto pour l'électronique grand public et le M2M aident les OEM grand public, les OEM industriels et les opérateurs de réseaux mobiles à surmonter ces problèmes de sécurité, notamment les risques de vol de propriété intellectuelle dus à l'environnement non réglementé dans lequel ces appareils évoluent.

7.2. Pilier numéro deux - La sécurisation du cloud :

Les menaces les plus pressantes proviennent de l'environnement d'entreprise ou du cloud auquel ces appareils sont connectés. Les solutions de Gemalto pour le chiffrement de données et la sécurité du cloud constituent un portefeuille complet à la disposition des prestataires de services cloud et des entreprises pour sécuriser leurs actifs. Notre solution d'attribution de licences et d'habilitation basée sur le cloud permet aux entreprises technologiques de déployer le plein potentiel de l'environnement cloud, sécurisant ainsi leur propriété intellectuelle.

7.3. Pilier numéro trois – La gestion du cycle de vie de la sécurité

Souvent ignorée, la gestion du cycle de vie de la sécurité des composants de sécurisation des appareils et du cloud constitue un élément clé d'une stratégie de sécurité numérique fiable et à long terme. La sécurité n'est pas une activité ponctuelle, mais un élément évolutif de l'écosystème de l'IoT. Ajout de nouveaux appareils, déclassement des appareils obsolètes, intégration des appareils à un nouvel écosystème cloud ou l'inverse, gestion des téléchargements de logiciels et micro logiciels sécurisés : l'ensemble de ces activités nécessitent une gestion complète des identités, des clés et des tokens.

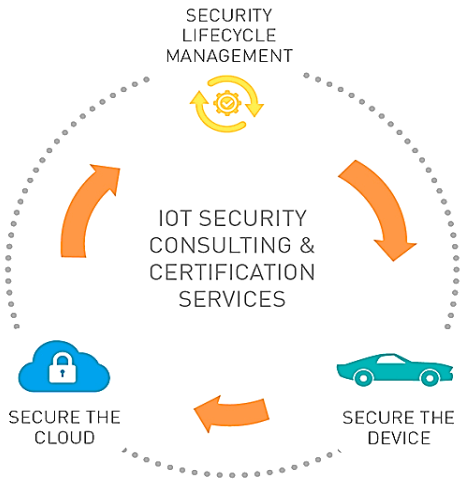

Figure 4.10 : piliers essentiels pour sécuriser un appareil IoT

Conclusion

En raison des capteurs et des positionneurs qui sont toujours allumés et en constante interaction avec leur corps et des autres dispositifs utilisés dans leur environnement, les utilisateurs auront plus de difficulté à séparer les différentes sphères de leur vie. Il devra également y avoir une responsabilité réelle pour les résultats des décisions que les soi-disant machines intelligentes prennent à notre sujet.

Dans un monde où nos activités et nos comportements quotidiens sont appelés à être de plus en plus mesurés, enregistrés et analysés, il est urgent que les concepteurs et les décideurs réfléchissent à la manière d'informer les consommateurs et les citoyens afin qu'ils sachent qui recueille leurs renseignements personnels, quels renseignements personnels sont recueillis, la manière dont ils sont conservés, utilisés et communiqués, à qui ils sont communiqués, et à quelles fins.

Si la transparence en ce qui concerne la collecte de données au moyen de dispositifs intelligents à l'ère de l'Internet des objets est importante pour nos relations avec le secteur privé, elle l'est tout autant pour nos relations avec le gouvernement. Il ne faudra guère nous surprendre que la manne de renseignements recueillis au moyen de l'Internet des objets à des fins commerciales suscite l'intérêt d'organismes d'application de la loi et de gouvernements.

Dans ce chapitre, nous avons présenté des perspectives motivées par l'évolution de l'Internet vers un Internet des objets. Nous avons montré que cette nouvelle technologie de rupture à enjeux socioéconomiques importants suscitera ses propres challenges de sécurité et de « privacy ». Afin d'apporter des réponses à ces problèmes de sécurité de l'Internet des objets, nous avons présenté un plan d'action à court, moyen et long terme. Ces trois phases porteront respectivement sur la sécurité efficace pour une informatique embarquée miniaturisée, la sécurité et privacy de l'informatique mobile omniprésente.

Chapitre 5

Chapitre 5: travaux de recherche et proposition

1. Introduction

L'Internet des objets (IoT, Internet of Things) confronte les entreprises du monde entier à des défis uniques en termes de sécurité, de confidentialité et de conformité. Contrairement à la technologie informatique traditionnelle où ces problèmes sont axés sur les logiciels et leur mode d'implémentation, l'IoT porte sur les effets de la convergence entre le monde informatique et le monde physique. La protection des solutions IoT implique un approvisionnement sécurisé des appareils, une connexion sécurisée entre ces appareils et le cloud et une protection efficace des données dans le cloud, dans le cadre du traitement et du stockage. Cependant, les appareils avec contraintes de ressources, la répartition géographique des déploiements et le grand nombre d'appareils inclus au sein d'une solution vont à l'encontre de ces fonctionnalités.

Ce chapitre se intéresse à des recherches proposée par les chercheurs et étudie de quelle façon notre proposition fournit une solution IoT cloud sécurisée et privée de bout en bout, intégrant la sécurité à chaque étape. À cet égard, l'approche (SDL) [1] représente la méthodologie de développement fondamentale. Elle est associée à une multitude de services de sécurité au niveau de l'infrastructure, comme le processus (OSA)[2], la Microsoft Digital Crimes Unit, (MSRC)[3] et le Centre de protection Microsoft contre les programmes malveillants.

[1] SDL : Security Development Lifecycle
[2] OSA: Operational Security Assurance
[3] MSRC : Microsoft Security Response Center

I. Partie 1 : Travaux de recherche

De nombreuses entreprises trouvent du marché de l'Internet des objets un champ fertile pour y investir. Intel, IBM et Google sont les trois entreprises principales dans le domaine. Chacune d'entre elles se contente de s'adapter rapidement à l'évolution de l'IoT en développant des solutions innovantes, basées sur les facilités *cloud*, pour la connectivité des objets à Internet, tout en assurant une bonne sécurité dans les différents niveaux de la connectivité allant des objets connectés jusqu'au cloud.

Sur la voie de la réalisation de la vision de l'internet des objets, des groupes de recherche réunissant des chercheurs académiques ou institutionnels se sont créés. Ce dans le but de développer des produits et des solutions avancées qui répondent aux besoins de l'IoT en tant que projet global qui vient de devenir une réalité, dont les avantages et les rendements seraient tout comme attendu. Dans ce contexte, des projets prometteurs se sont déjà lancés. ERCIT (European Research Cluster on the Internet of Things). Représente un large éventail de projets de recherches concernant l'application de l'Internet des objets avec des dimensions européennes. Garantir la collaboration et la communication entre ces projets est un prérequis essentiel pour une industrie compétitive et un déploiement sécurisé et sûr de l'IoTen Europe. **[web 40]**

1. Commercialisation et projets de recherche

1. **Butler est un projet européen**, son but est le développement des applications sécurisées et intelligentes basées sur des systèmes d'information omniprésents et contextuels. Bulter s'intéresse aux scénarios du genre villes intelligentes, maisons intelligentes, applications de santé assistées par l'informatique ubiquitaire et des applications commerciales intelligentes. En ce qui concerne les exigences de sécurité, le projet vise à permettre aux utilisateurs de gérer leurs profils distribués, ce qui implique le contrôle de duplication de données et des identités utilisées par les applications distribuées. L'objectif final étant de mettre en œuvre un système capable d'intégrer des données dynamiques de l'utilisateur (par exemple, la localisation, le comportement) dans les protocoles de sécurité. **[web 37].**
2. **Le projet *Hydra* cofinancé par la commission européenne**, sert à développer une couche middleware pour la connexion des réseaux d'objets intelligents à Internet, en se basant sur une architecture orientée services (SOA : Service Oriented Architecture). Ce projet envisage les questions de sécurité distribuée et de confiance sociale. Une telle middleware permet aux développeurs d'incorporer des dispositifs matériellement hétérogènes dans leurs applications en offrant des interfaces de service Web facile à utiliser pour contrôler tout type de périphérique sans se soucier aux différentes technologies de transmission adoptées dans le réseau, telles que Bluetooth, ZigBee et Wifi. Hydra incorpore des mécanismes pour la découverte des dispositifs et des services, une architecture orientée modèle sémantique et même les communications P2P (Peer to Peer). **[web 38],**

3. **Un projet de recherche prometteur dénommé NITRD** (Networking and Information Technology Research and Development), commencé en 2012. Ce projet regroupe une dizaine d'agences fédérales, telles que NASA (National Aeronautics and Space Administration) et DARPA (Defense Advanced ResearchProjects Agency). L'objectif est de développer des infrastructures intelligentes pour la concrétisation efficace des différents scénarios applicatifs de l'IoT. **[web 39]**

2. La sécurité pour les réseaux de capteurs

Les exigences de sécurité, dans les réseaux de capteurs, dépendent de la nature de l'application. Les applications militaires sont très exigeantes en sécurité, voire non tolérante à ce niveau ; car le réseau est une épée a doublé tranches, et peut devenir une arme ennemie.

Mais dans des applications de surveillance de l'environnement, par exemple, la Sécurité n'est pas très exigeante. L'application d'un simple mécanisme demeure suffisante pour protéger le réseau des attaques primaires.

Parmi les attaque malveillante on trouve L'attaque Sybille qui est tout d'abord a été introduite et décrite par Douceur dans le contexte des réseaux pair a pair **[77].** Dans cette attaque, un nœud malveillant peut revendiquer différentes identités dans le but de prendre de l'avantage sur les nœuds légitimes. Dans **[79],** Newsome et al. A introduit une taxonomie dans laquelle ils proposent de classer les attaques Sybille selon les trois dimensions suivant:

- communication directe vs communication indirecte
- identités fabriquées vs identités volées
- simultanéité

2.1 Travaux des recherches

2.1.1 Certification approuvée

Selon Douceur **[77],** dans un environnement de calcul distribue, un attaquant peut de façon illégitime présenté des identités multiples en se servant de la même entité physique. En revendiquant ces différentes identités, l'entité malveillante est avantagée par rapport aux autres entités. Cela pose une menace de sécurité, particulièrement dans les systèmes pair a pair qui mettent en œuvre la réplication de contenu, ou dans les schémas de fragmentation sur plusieurs pairs pour assurer de la disponibilité et de la sécurité et qui reposent sur l'existence de pairs indépendants avec des identités différentes.

Dans **[77],** Douceur affirme qu'a moins qu'une certification approuvée soit mise en œuvre, les systèmes pair a pair seraient exposes aux attaques Sybille. Dans la certification approuvée, une autorité centrale est utilisée dans le but de se porter garante de la correspondance un a un entre chaque entité et chaque identité. Cependant, la certification approuvée repose sur une autorité Centrale qui doit garantir que chaque entité est affectée a exactement une seule identité. Dans ce papier, Douceur n'offre aucune méthode pour assurer une telle unicité.

Newsome et al. Étudient dans **[79]** les attaques Sybille dans le contexte des réseaux de capteurs et définissent une attaque Sybille comme un capteur malveillant qui prend illégitimement de multiples identités. Les identités additionnelles sont considérées comme étant des nœuds Sybille. Ces nœuds Sybille peuvent soit communiquer directement ou indirectement via un nœud malveillant. Ils peuvent obtenir des identités soit volées soit fabriquées. Ces identités peuvent participer soit simultanément dans le réseau soit de façon non simultanée.

Dans **[79],** Newsome et al. Citent aussi la certification approuvée comme un moyen d'éliminer les attaques Sybille. Cette approche est la même que celle proposée dans **[77].** La seule différence est que dans **[79],** Newsome et al. On étudie les attaques Sybille dans le contexte des réseaux de capteurs. En effet, dans les réseaux de capteurs, il peut y avoir une autorité centrale qui gère le réseau et donc qui connait toutes les identités des nœuds déployés. Ainsi, l'autorité centrale peut

détecter des attaques Sybille en interrogent le réseau et en comparant les résultats des requêtes avec le déploiement connu.

Bien que la certification approuvée soit l'approche la plus citée dans la littérature pour lutter contre les attaques Sybille **[78]**, elle présente les désavantages suivants :

- La liste des identités connue doit être protégée d'une éventuelle modification malicieuse. En effet, si l'adversaire est capable d'ajouter de nouvelles identités a cette liste, il pourra être capable d'ajouter des nœuds Sybille au réseau.
- L'entité qui gère le réseau de capteur doit de façon sécurisée être capable de rajouter de nouveaux nœuds au réseau et doit maintenir et demander a connaitre des informations sur la topologie actuellement connue.

Elle engendre une grande surcharge due aux nombreuses communications requises entre l'autorité centrale et les autres nœuds du réseau.

2.1.2 Test de ressources

Le but du test de ressources est d'essayer de déterminer si certaines identités possèdent moins de ressources qu'elles sont supposées avoir si elles étaient indépendantes.

Dans **[77],** Douceur définit le modèle de l'attaquant comme étant un environnement générique de calcul distribue qui manque d'autorité centrale. Dans ce modèle, les entités communiquent via des messages broadcastes. Certaines entités sont honnêtes, d'autres non. Une entité honnête présentera une seule identité légitime alors qu'une malhonnête peut essayer de présenter non seulement l'identité légitime mais aussi une ou plusieurs identités illégitimes. Les messages sont supposés être reçus de la part de toutes les entités et les ressources (mémoire, communication, calcul) de chaque entité physique (qu'elle soit honnête ou non) sont supposées être limitées.

Pour accepter une identité, un vérifieur peut lancer un défi exigeant en ressources en broadcastant une requête aux identités et valide seulement les identités dont les réponses ont lieu dans un intervalle de temps donne. Ensuite, les entités peuvent mettre en commun les identités qu'elles ont validées séparément.

Pour prouver qu'une telle solution est impraticable, Douceur souligne que cette solution peut être possible seulement sous des hypothèses irréalistes de ressources et de coordination parmi les entités. En effet, toutes les entités doivent avoir approximativement les mêmes contraintes de ressources et les identités doivent être validées simultanément par toutes les entités coopérants.

Dans **[79],** Newcomen et al. Proposent aussi l'utilisation du test de ressources.

Dans cette approche, ils supposent que tout capteur physique possède seulement une seule radio et que cette radio est incapable d'envoyer et de recevoir simultanément sur plus d'un canal. Dans ce papier, les auteurs considèrent que chaque vérifieur attribue un canal différent à chacun de ses n voisins afin d'y broadcaster un message. Ensuite, il choisit aléatoirement un canal sur lequel il décide d'écouter. Si le vérifieur arrive à attendre le message sur ce canal, cela signifie que le voisin a qui était attribué ce canal est légitime. Dans le cas échéant, cela signifie que ce voisin est un nœud Sybille. La difficulté de cette approche est que nous n'avons pas toujours autant de canaux que de voisins. Dans ce cas, si nous n'avons pas suffisamment de canaux à attribuer, cela peut prendre beaucoup de temps pour détecter une attaque Sybille. Elle peut même demeurer in détectée si il y a autant ou plus de nœuds Sybille que de canaux. Cependant, nous ne pouvons malheureusement pas envisager d'utiliser une telle approche vu qu'elle ne peut malheureusement convenir a des systèmes

distribues dans un réseau de large étendue. En effet, la validation peut nécessiter d'importants couts en termes d'énergie

II. Partie 2 : Proposition

1. Introduction

Notre idée proposée offre des fonctionnalités uniques, assurant la simplicité, la transparence et, surtout, la sécurité de l'approvisionnement des appareils IoT, de la connexion à ces appareils et du stockage des données qu'ils fournissent. Dans ce qui suit, nous examinons les fonctionnalités de sécurité de notre plateforme et les stratégies de déploiement garantissant la résolution des problèmes de sécurité, de confidentialité et de conformité.

2. La sécurité commence par un modèle de menace

2.1. Présentation

L'objectif de la modélisation des menaces est de comprendre comment une personne malveillante peut être en mesure de compromettre un système avant de s'assurer que les préventions appropriées sont en place. La modélisation des menaces impose à l'équipe de conception de tenir compte des préventions lors de la conception du système plutôt qu'à l'issue de son déploiement. Ce point est extrêmement important, car dans la pratique modifier des solutions de défense de la sécurité sur une multitude de périphériques est infaisable, sujet aux erreurs et rend les clients vulnérables.

De nombreuses équipes de développement font un excellent travail en prenant en compte les exigences fonctionnelles du système dont bénéficient les clients. Néanmoins, l'identification des modes non évidents de compromission du système par une personne malveillante se révèle plus difficile. La modélisation des menaces peut aider les équipes de développement à comprendre ce qu'une personne malveillante peut effectuer et pourquoi. Cette modélisation est un processus structuré qui crée une discussion sur les décisions en matière de conception de la sécurité dans le système, et qui apporte des modifications tout au long de la conception qui ont des incidences sur la sécurité. Si un modèle de menace consiste tout simplement en un document, celui-ci représente un moyen idéal de s'assurer de la continuité des connaissances et de la conservation des leçons retenues. Par ailleurs, il contribue à l'intégration rapide des nouvelles équipes. Enfin, la modélisation des menaces vous permet de prendre en considération d'autres aspects de la sécurité,

2.2. Quand mettre en place un modèle de menace

La modélisation des menaces apporte davantage de valeur si elle est incorporée dans la phase de conception. Lors de la conception, vous bénéficiez d'une plus grande flexibilité pour apporter les modifications qui élimineront les menaces. L'élimination des menaces au cours de la conception est en effet le résultat souhaité. C'est en effet bien plus facile que d'ajouter des préventions, les tester et s'assurer qu'elles restent en cours. Par ailleurs, procéder à de telles éliminations n'est pas toujours possible. Il est plus difficile d'éliminer des menaces lorsqu'un

produit progresse dans son cycle de vie, et cela nécessite plus de travail et des compromis beaucoup plus difficiles que la mise en place d'une modélisation des espaces au début du développement.

2.3. Quels éléments inclure dans le modèle de menace

Vous devez inclure la solution dans son ensemble dans le modèle de menace et vous concentrer sur les domaines suivants :

- Les fonctionnalités de sécurité et de confidentialité
- Les fonctionnalités dont les échecs relèvent de la sécurité
- Les fonctionnalités qui touchent une délimitation d'approbation

de développement à comprendre ce qu'une personne malveillante peut effectuer et pourquoi.

2.4. Procédure à suivre pour modéliser les menaces

Le processus de modélisation des menaces est composé des quatre étapes suivantes :

- Modéliser l'application
- Énumérer les menaces
- Prévenir les menaces
- Valider les préventions

2.5. Étapes du processus

Vous devez garder à l'esprit trois règles élémentaires lorsque vous créez un modèle de menace:

1. Créez un diagramme de l'architecture de référence.
2. Commencez par la logique de largeur. Obtenez d'abord une vue d'ensemble du système pour le comprendre avant de l'explorer en profondeur. Cette façon de procéder garantit que vous explorez en profondeur les bons emplacements.
   C'est vous qui conduisez le processus, et non l'inverse. Si vous rencontrez un problème lors de la phase de modélisation et souhaitez l'examiner.

3 . Plateforme proposée :

3.1. Approvisionnement et authentification sécurisés des appareils

Notre plateforme –baptisée **PFIOTSecurity**- propose de sécuriser les appareils emportés sur le terrain en fournissant une clé d'identité unique pour chacun d'eux. Cette clé peut être utilisée par l'infrastructure IoT pour communiquer avec l'appareil, lorsqu'il est en cours d'utilisation. La configuration du processus est simple et rapide. La clé générée avec un ID

d'appareil sélectionné par l'utilisateur constitue la base d'un jeton utilisé pour toutes les communications entre l'appareil et la plateforme.

On peut suggérer que Les ID d'appareil peuvent être associés à un appareil lors de la fabrication (flashés dans un module matériel de confiance) ou utiliser une identité fixe existante en tant que proxy (par exemple, les numéros de série de processeur). Il est complexe de modifier ces informations d'identification dans l'appareil. Il convient donc d'introduire des ID d'appareil logique, au cas où le matériel sous-jacent subit des modifications et où l'appareil logique reste le même. Dans certains cas, l'association d'une identité d'appareil peut se produire au moment du déploiement de l'appareil (par exemple, un ingénieur de terrain authentifié configure physiquement un nouvel appareil en communiquant avec le serveur principal de solution). Le registre des identités de la plateforme stocke de manière sécurisée les identités des appareils et les clés de sécurité d'une solution. Des identités d'appareil peuvent être ajoutées à une liste verte , individuellement ou en groupe, permettant un contrôle total de l'accès à l'appareil.

Les stratégies de contrôle d'accès de notre plateforme dans le cloud permettent d'activer et de désactiver l'identité de tout appareil, offrant le moyen de dissocier un appareil d'un déploiement IoT si nécessaire. Cette association et cette dissociation d'appareils sont basées sur l'identité de chaque appareil. La plateforme proposée doit assurer une détection des intrusions et une prévention continue, une prévention contre les attaques de service et des tests d'intrusion réguliers. Elle doit offrir également des outils d'investigation qui permettent d'identifier et d'atténuer les menaces. L'authentification multi-facteur fournit une couche de sécurité supplémentaire pour l'accès au réseau par les utilisateurs finaux.

Pour présenter clairement les fonctionnalités de sécurité et de confidentialité intégrées à **PFIOTSecurity**, nous avons fractionné la suite selon les trois principaux domaines de sécurité.

La sécurité de l'authentification multi-facteur repose sur son approche en couche. Compromettre plusieurs facteurs d'authentification présente un défi de taille pour les attaquants. Même si un attaquant réussit à connaître le mot de passe de l'utilisateur, ce dernier est inutile sans posséder l'appareil de confiance. Si l'utilisateur perd l'appareil, la personne qui l'a trouvé ne pourra pas l'utiliser sans connaître le mot de passe de l'utilisateur. Lorsqu'un utilisateur se connecte, une vérification supplémentaire lui est envoyée. Voici la liste des méthodes qui peuvent être utilisées pour cette seconde vérification.

Méthode de vérification	Description
Appel téléphonique	Un appel est passé sur le téléphone enregistré de l'utilisateur demandant de vérifier qu'il se connecte en appuyant sur le symbole # ou en entrant un code confidentiel.
SMS	Un SMS sera envoyé sur le téléphone de l'utilisateur avec un code à six chiffres. Entrez ce code pour finaliser le processus de vérification.
Code de vérification de l'application mobile	L'application mobile, qui est exécutée sur le smartphone d'un utilisateur, affiche un code de vérification à 6 chiffres qui change toutes les 30 secondes. L'utilisateur trouve le code le plus récent et l'entre dans la page de connexion pour terminer le processus de vérification. Cela se produit si vous avez sélectionné un code de vérification comme méthode de vérification principale.

Tableau 5.1 : Approvisionnement et authentification sécurisés des appareils

D'autres fonctionnalités de sécurité des appareils sont disponibles :

- Les appareils n'acceptent pas les connexions réseau non sollicitées. Ils établissent tous les itinéraires et connexions pour le trafic sortant uniquement. Pour qu'un appareil reçoive une commande de la part d'un serveur principal, il doit établir une connexion pour vérifier toutes les commandes en attente de traitement. Lorsqu'une connexion entre l'appareil et **PFIOTSecurity** est établie en toute sécurité, l'échange de messages entre le cloud et l'appareil peut se faire en toute transparence.

- Les appareils se connectent ou établissent des itinéraires vers des services bien connus auxquels ils sont couplés.
- L'autorisation et l'authentification au niveau du système utilisent des identités par appareil. Les informations d'identification et autorisations d'accès sont ainsi révocables quasi instantanément.

3.2. Vue d'ensemble de PFIOTSecurity

PFIOTSecurity est un service entièrement géré qui permet des communications bidirectionnelles fiables et sécurisées entre des millions d'appareils IoT et un serveur principal de solution. Cette plateforme :

- Fournit plusieurs options de communication appareil vers cloud et cloud vers appareil, y compris la messagerie unidirectionnelle, le transfert de fichiers et les méthodes de demande-réponse.
- Intègre une fonctionnalité de routage des messages déclaratifs vers d'autres services.
- Fournit un stockage utilisable dans une requête pour les métadonnées d'appareil et les informations d'état synchronisées.
- Assure la sécurité des communications et le contrôle d'accès grâce aux clés de sécurité par appareil ou aux certificats X.509[4].
- Fournit une surveillance complète de la connectivité des appareils et des événements de gestion de l'identité des appareils.
- Inclut des bibliothèques d'appareils pour les langages et les plateformes les plus courants

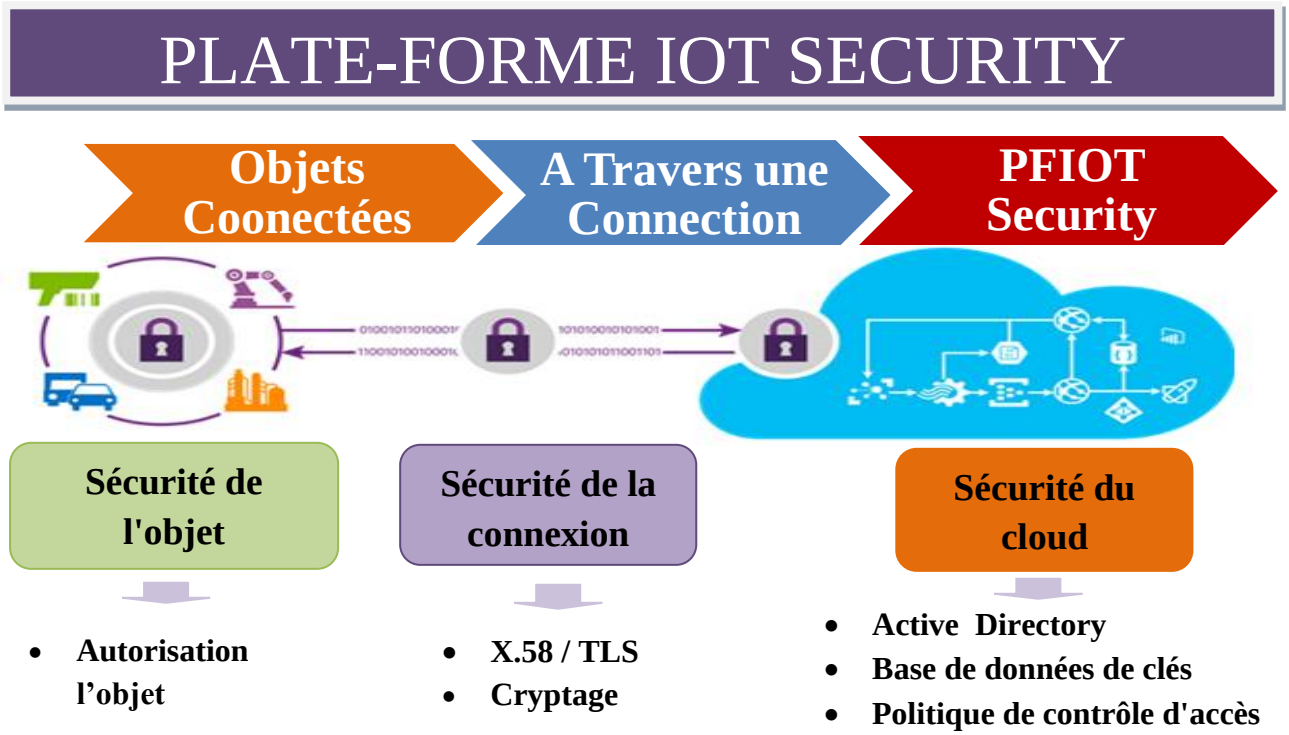

Figure 5.1 : Plateforme IOT sécurisée proposée

3.3. Comment PFIOTSecurity fonctionne-t-il ?

Notre plateforme met en œuvre un modèle de communications pour gérer les interactions entre les appareils et un serveur principal de service. Ce modèle établit les principes suivants :

- La sécurité est prioritaire sur toutes les autres fonctionnalités.

Les appareils n'acceptent pas les informations réseau non sollicitées. Un appareil établit tous les itinéraires et connexions pour le trafic sortant uniquement. Pour qu'un appareil reçoive une

[4]**X.509** est une norme spécifiant les formats pour les certificats à clé publique, les listes de révocation de certificat, les attributs de certificat, et un algorithme de validation du chemin de certification, définie par l'Union internationale des télécommunications (UIT)[1]. X.509 établit entre autres un format standard de certificat électronique et un algorithme pour la validation de chemin de certification

commande de la part du serveur principal de la solution, il doit régulièrement établir une connexion pour vérifier toutes les commandes en attente de traitement.

- Les appareils doivent uniquement connecter ou établir des itinéraires vers des services bien connus auxquels ils sont couplés.
- L'itinéraire de communication entre l'appareil et le service ou entre l'appareil et la passerelle est sécurisé au niveau du protocole d'application.
- L'authentification et l'autorisation au niveau du système sont basées sur les identités par appareil. Les autorisations et informations d'identification sont ainsi révocables presque instantanément.
- La communication bidirectionnelle des appareils connectés de façon sporadique en raison de problèmes d'alimentation ou de connectivité peut être facilitée par la mise en attente de commandes et de notifications aux appareils jusqu'à ce que qu'un appareil soit connecté pour les recevoir. Notre plateforme gère des files d'attente spécifiques à un appareil pour les commandes qu'il envoie.

3.4. Une connectivité sécurisée

La durabilité de la messagerie est une fonctionnalité importante de toute solution IoT. La nécessité de fournir des commandes et/ou de recevoir des données à partir des appareils de façon durable est d'autant plus importante que les appareils IoT sont connectés par le biais d'Internet ou d'autres réseaux similaires, qui ne sont pas toujours fiables. **PFIOTSecurity** assure la durabilité de la messagerie entre le cloud et les appareils par le biais d'un système d'accusés de réception en réponse aux messages. Cette durabilité peut être renforcée en mettant en cache les messages dans **PFIOTSecurity** pendant sept jours maximum pour la télémétrie et deux jours maximum pour les commandes (comme exemple).

L'efficacité est essentielle pour assurer la conservation des ressources et le fonctionnement dans un environnement avec contraintes de ressources. Pris en charge par **PFIOTSecurity**, le protocole HTTPS (HTTP Secure), version sécurisée standard du protocole populaire http, permet une communication efficace. Les protocoles AMQP (Advanced Message Queuing Protocol) et MQTT (Message Queuing Telemetry Transport), pris en charge, sont conçus non seulement pour leur efficacité en termes d'utilisation des ressources, mais également pour leur fiabilité en matière de remise des messages.

L'évolutivité implique la possibilité d'interagir en toute sécurité avec un large éventail d'appareils. **PFIOTSecurity**assure une connexion sécurisée avec les appareils IP et non-IP. Les appareils IP peuvent se connecter et communiquer directement avec la plateforme par le biais d'une connexion sécurisée. Les appareils non-IP présentent des contraintes de ressources et se connectent uniquement avec des protocoles de communication de courte distance, tels que Zwave, ZigBee et Bluetooth. Une passerelle de champ est utilisée pour agréger ces appareils. Elle effectue la traduction de protocole pour établir une communication bidirectionnelle sécurisée avec le cloud.

D'autres fonctionnalités de sécurité des connexions sont disponibles :

- L'itinéraire de communication entre les appareils et la plateforme, ou entre les passerelles et la plateforme, est sécurisé à l'aide du protocole standard TLS (Transport Layer Security)[5].

- Afin de protéger les appareils contre les connexions entrantes non sollicitées, la plateforme n'ouvre aucune connexion vers l'appareil. L'appareil lance toutes les connexions.

Notre plateforme stocke durablement les messages pour les appareils et attend leur connexion. Ces commandes sont stockées pendant deux jours. Les appareils confrontés à des problèmes de connectivité ou d'alimentation peuvent ainsi se connecter de façon sporadique pour recevoir ces commandes. Notre plateforme gère une file d'attente spécifique pour chaque appareil.

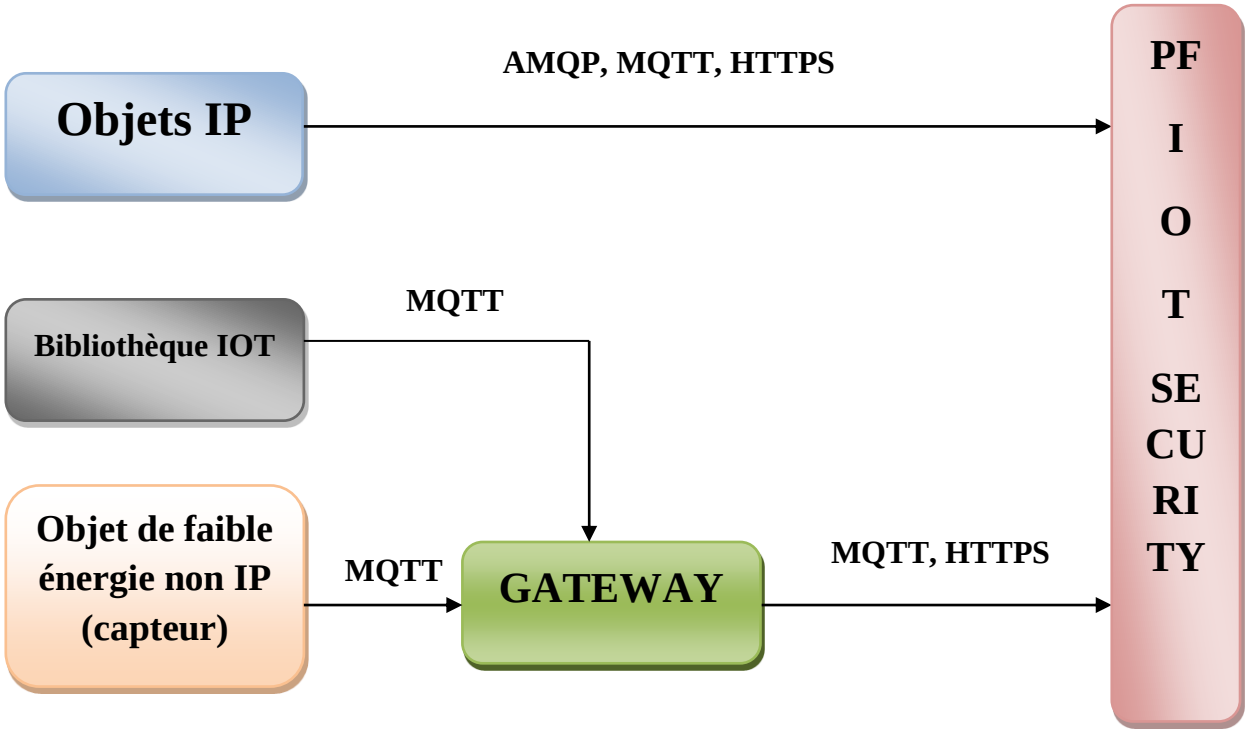

Figure 5.2 : Une connectivité sécurisée

Une fois les données dans le cloud, elles peuvent être traitées et stockées dans n'importe quel flux de travail défini par l'utilisateur. L'accès à chaque partie des données est contrôlé avec Active Directory, en fonction du service de stockage utilisé.

Toutes les clés utilisées par l'infrastructure IoT sont stockées dans le cloud de manière sécurisée. Les clés peuvent être substituées si un réapprovisionnement est nécessaire. Les données peuvent être stockées dans des **bases de données SQL**, ce qui permet de définir le niveau de sécurité souhaité. En outre, **PFIOTSecurity** fournit un moyen de surveiller et d'auditer tous les accès à vos données pour vous informer de toute intrusion ou de tout accès non autorisé.

[5]Transport Layer Security (TLS), et son prédécesseur Secure Sockets Layer (SSL), sont des protocoles de sécurisation des échanges sur Internet. Le protocole SSL a été développé à l'origine par Netscape. L'IETF, en a poursuivi le développement en le rebaptisant Transport Layer Security (TLS). On parle parfois de SSL/TLS pour désigner indifféremment SSL ou TLS.

4. Conclusion

L'Internet des objets concerne avant tout vos propres activités, les choses essentielles pour l'entreprise. L'IoT peut apporter une formidable valeur ajoutée à une entreprise, en lui permettant de réduire ses coûts, d'augmenter son chiffre d'affaires et de transformer son activité. La réussite de cette transformation repose en grande partie sur le choix du fournisseur de logiciels et de services IoT approprié. Il s'agit de trouver un fournisseur qui va non seulement déclencher cette transformation en comprenant les besoins et les exigences de l'entreprise, mais également fournir des services et des logiciels intégrant la sécurité, la confidentialité, la transparence et la conformité comme éléments de conception essentiels. Fort d'une expérience incomparable en développement et déploiement de services et logiciels sécurisés, Microsoft continue de s'imposer comme véritable leader dans cette nouvelle ère de l'Internet des objets.

PFIOTSecurity intègre des mesures de sécurité par conception, assurant la surveillance sécurisée des ressources pour améliorer l'efficacité, favorisant les performances opérationnelles pour permettre l'innovation et employant des analyses de données avancées pour transformer les activités. Grâce à son approche en couche de la sécurité, à sa multitude de fonctionnalités de sécurité et à ses modèles de conception, notre approche permet de déployer une infrastructure de confiance, pouvant transformer toute activité.

Conclusion générale

À l'origine de ce travail de ce mémoire, nous avons posé plusieurs problématiques pour la réalisation de l'Internet des objets en tant que technologie capable de concrétiser à grande échelle divers scénarios relatifs à l'informatique ubiquitaire et à l'intelligence ambiante.

Parmi ces problématiques, nous considérons spécifiquement, d'une part, celles liées à l'architecture, au développement, au déploiement et à la sécurité pour l'internet des objets.

Maisons connectées, habits intelligents, pèse-personnes communicants…. La croissance exponentielle du nombre d'objets connectés est le témoin d'une tendance particulièrement impactant dans le monde numérique, portée par les usages du grand public. Les objets connectés apportent une forme de maîtrise de l'environnement des utilisateurs, que ce soit sur les aspects de santé, de gestion de l'énergie ou la maîtrise de l'information.

Cette tendance technologique innovante va également avoir un impact fort pour l'entreprise, que ce soit au niveau de son modèle d'affaires, de ses processus internes ou des nouveaux risques associés. En effet, l'Internet des Objets adresse l'ensemble des tendances structurantes de la transition numérique : primauté de l'expérience client, organisation et management, ressources et flux.

La question qui se pose alors est de transformer ces objets en valeur, c'est-à-dire en de nouveaux services qui vont avoir un impact significatif dans la vie des citoyens, des employés, de l'écosystème.

A l'image du Big Data auquel il est fortement lié, l'Internet des Objets offre de vastes possibilités pour l'évolution des modèles d'affaires des entreprises vers le numérique, que ce soit dans la proposition de nouveaux services aux clients et usagers, dans « l'augmentation »de la relation client ou encore dans la valorisation des données récupérées.

L'Internet des Objets n'en est qu'à ses balbutiements, mais il est impossible d'ignorer l'importance que cette tendance va prendre dans la société et dans nos entreprises. Il est alors primordial de se pencher dès aujourd'hui sur cette problématique afin de développer l'écosystème qui favorisera et encadrera son intégration dans notre quotidien. La collaboration de tous sera nécessaire pour mettre en place les infrastructures indispensables, et développer une offre de confiance numérique.

La véritable problématique de l'Internet des Objets est bien ici, dans cette valeur de la culture numérique qu'est la confiance, car dans ce contexte, les utilisateurs confient leur sin formations à des outils dont la chaîne de gestion sera partagée par un grand nombre d'acteurs, privés et publics.

L'entreprise qui aura mis en place des nouveaux modèles d'affaire liés à l'Internet des Objets tout en améliorant ses processus internes devra donc bien s'ouvrir à la co-création de valeur avec

l'ensemble de son écosystème. Cela ouvre d'ailleurs certaines questions qui ne trouvent pas encore de réponses :

- Comment gérer les objets en fin de vie (aspect Green) ?
- Quelle sera l'acceptation du public sur les problématiques de traçage ?
- Les nouvelles générations vont-elles avoir le même rapport à la notion de vie privée ?

L'Internet des Objets n'est pas qu'une problématique technologique (sur laquelle on sait déjà être innovant), mais également une problématique sociétale et éthique. Charge à nous de récolter les fruits de ces innovations prometteuses pour l'entreprise dans une démarche constructive et responsable.

RÉFÉRENCES

[1] [Benjamin Renaut 13-14], "Hadoop/Big Data", Université De Nice Sophia-Antipolis, 114p.

[2] [Mathieu Millet 13], "Big Data".

[3] [Angeline Kone 13] "Big Data".

[4] [Pierre Nerzic Mars 16] "Outils Hadoop Pour Le BigData" .

[5] [Mickal Baron 14] "Généralités Sur Hdfs Et MapReduce".

[6] [Mekideche Mounir Avr 15] "Conception Et Implémentation D'un Moteur De Recherche A Base D'une Architecture Hadoop (Big Data)".

[7] [Mickael Baron Avr 14] "Tutoriel D'introduction A Apache Hadoop" .

[8] [Bernard Espinasse Avr 13] "Introduction Aux Systèmes NOSQL (Not OnlySql)", Ecole Polytechnique Universitaire De Marseille, 19p.

[9] [Matteo Di Maglie Sept 12] "Adoption D'une Solution NOSQL Dans L'entreprise", Travail De Bachelor Réalisé En Vue De L'obtention Du BachelorHes, Carouge.

[10] [Khaled Tannir Dec 15] " NOSQL (Montréal -01 Décembre 2015)".

[11] [Qi Zhang • Lu Cheng • Raouf Boutaba 10] " Cloud computing: state-of-the-art and research challenges".

[12] [Hüsemann, S 10]"Systèmes d'information. Fribourg".

[13] [Krutz, R. L., & Dean Vines,. 10] "Cloud Security: A Comprehensive Guide to Secure Cloud Computing. Wiley. ".

[14] [Hurwitz, J., Bloor, R., Kaufman, M., & Halper, F... 10] "Cloud Computing for Dummies. Wiley. ".

[15] [Hammou FADILI] "WEB 2.0 & WEB 3.0 popularisation de la création et de la promotionculturelle et scientifique"

[16] [Thibaut Watrigant 14] "Comment communiquent vos objets connectés" .

[17] [P. Karn, MACA, A 1990] "New Channel Access Protocol for Packet Radio, ARRL/CRRL, 9th Computer Networking Conference, pp 134-140, ".

[18] [N. Abramson 1985] "Development of the ALOHANET, IEEE Transactions on Information Theory, vol IT-31, pp 119-123,"

[19] National Intelligence Council, Disruptive Civil Technologies — Six Technologies withPotential Impacts on US Interests Out to 2025—Conference Report CR 2008–07, April 2008.

[20] [ITSEC 91] ITSEC. Information Technology Security Evaluation Criteria – Provisional Harmonised Criteria. Commission of the European Communities. DG XIII., Document COM (90) 314, Office for Official Publications of the European Communities, Luxembourg, juin 1991. ISBN-10 9-2826-3004-8. http://www.ssi.gouv.fr/site_documents/ ITSEC/ITSEC-uk.pdf.

[21] [Common Criteria 12] Common Criteria. Common Criteria for Information Technology Security Evaluation - Part 1: Introduction and general model. Version 3.1, révision 4, numéro CCMB-2012-09-001, septembre 2012. http://www.commoncriteriaportal.org/files/ccfiles/CCPART1V3.1R4.pdf.

[22] [Laprie 04] Jean-Claude Laprie. "Sûreté de fonctionnement des systèmes : concepts de base et Terminologie. Revue de l'Électricité et de l'Électronique (REE), (11):95–105, " novembre 2004.

[23] [Laprie et al. 96] Jean-Claude Laprie, Jean Arlat, Jean-Paul Blanquart, Alain Costes, Yves Crouzet, Yves Deswarte, Jean-Charles Fabre, Hubert Guillermain, Mohamed Kaâniche, Karama Kanoun, Corinne Mazet, David Powell, Christophe Rabéjac, et Pascale Thévenod. Guide de la Sûreté de Fonctionnement. Cépaduès, 2e dition, 1996. ISBN 978-2-854-28382-2.

[24] [Adelsbach et al. 03] André Adelsbach, Dominique Alessandri, Christian Cachin, SadieCreese, Yves Deswarte, KlauseKursawe, Jean-Claude Laprie, David Powell, Brian Randell, James Riordan, Peter Ryan, William Simmonds, Rober Stroud, Paulo Veríssimo, Michael Waidner, ET Andreas Wespi. MAFTIA, Malicious and Accidental-Fault Tolerance for Internet Applications: Conceptual Model and Architecture. David Powell ET Robert Stroud, éditeurs. Rapport technique, 31 janvier 2003. Research Project IST-1999-11583, delivrable D21. http://research.cs.ncl.ac.uk/cabernet/www.laas.research.ec.org/maftia/deliverables/D21.pdf

[25] [Anderson 80] James P. Anderson. Computer security threat monitoring and surveillance. Rapport technique, Contrat numéro 79F296400, James P. Anderson Co., Washington (WA, USA), 26 février 1980. http://csrc.nist.gov/publications/history/ande80.pdf.

[26] [Parker 75] Donn B. Parker. Computer Abuse Perpetrators and Vulnerabilities of Computer Systems. In Proceedings of the National Computer Conference and Exposition, American Federation of Information Processing Societies (AFIPS). ACM Press, New York (NY, USA), 7-10 juin 1975. http://dx.doi.org/10.1145/1499799.1499810.

[27] [Shirey 94] RobertW. Shirey. Security Architecture for Internet Protocols: A Guide for Protocol Designs and Standards. novembre 1994. Internet Draft: draft-irtf-psrg-secarch-sect1-00.

[28] [Abbott et al. 76] Robert P. Abbott, Janet S. Chin, James E. Donnelley, William L. Konigsford, Shigeru Tokubo, ET Douglas A.Webb. Security Analysis and Enhancements of Computer Operating Systems. Rapport technique numéro NBSIR 76-1041, Institute for Computer Sciences and Technology, National Bureau of Standards, Washington DC, (WA, USA) avril 1976. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA436876.

[29] [Cheheyl et al. 81] Harris M. Cheheyl, Morrie Gasser, George A. Huff, ET Jonathan K. Millen. Verifying Security. ACM Computing Surveys (CSUR), 13(3):279–339, ACM Press, New York (NY, USA), septembre 1981. http://doi.acm.org/10.1145/356850.356853.

[30] [Fagan 76] Michael E. Fagan. Design and Code Inspections to Reduce Errors in Program Development. IBM Systems Journal, 15(3):182–211, IBM Corporation, Riverton (NJ, USA), juin 1976. ISSN 0018-8670. http://dx.doi.org/10.1147/sj.382.0258.122

[31] [Dyer 92] Michael Dyer. The Cleanroom Approach to Quality Software Development. John Wiley & Sons, Inc., New York (NY, USA), 1992. ISBN 979-0-471-54823-5.

[32] [van Emden 92] Maarten H van Emden. Structured Inspections of Code. Journal of Software Testing, Verification, and Reliability, 2:133–153, John Wiley & Sons, Inc., New York (NY, USA), 1992. http://dx.doi.org/10.1002/stvr.4370020304.

[33] [Ebenau et Strauss 94] Robert G. Ebenau et Susan H. Strauss. Software Inspection Process. McGraw-Hill systems design & implementation series. McGraw-Hill, New York (NY, USA), 1994. ISBN 978-0-070-62166-4.

[34] [Myers et al. 04] Glenford J. Myers, Corey Sandler, Tom Badgett, ET Todd M. Thomas. The Art of Software Testing. Business Data Processing: a Wiley Series. John Wiley & Sons, Inc., New York (NY, USA), 2004. ISBN 978-0-471-46912-4. http://dx.doi.org/10.1002/stvr.322/.

[35] [Cousot et Cousot 77] Patrick Cousot et RadhiaCousot. Abstract interpretation: a unified lattice Model for static analysis of programs by construction or approximation of fixpoints. In Conference Record of the Fourth Annual ACM SIGPLAN-SIGACT Symposium 120

On Principles of Programming Languages, pages 238–252, Los Angeles (CA, USA). ACM Press (NY, USA), 1977.

[36] [Cousot 01] Patrick Cousot. Abstract Interpretation Based Formal Methods and Future Challenges. In Reinhard Wilhelm, éditeur : Informatics - 10 Years Back. 10 Years Ahead., Volume 2000 de Lecture Notes in Computer Science, pages 138–156. Springer Berlin Heidelberg, 2001. ISBN 978-3-540-41635-7. http://dx.doi.org/10.1007/3-540-44577-3_10.

[37] [Floyd 67] Robert W. Floyd. Assigning Meanings to Programs. Mathematical Aspects of Computer Science, 19(19-32):19–32, American Mathematical Society, Providence (RI, USA), 1967. ISBN 978-0-821-86728-0. http://www.eecs.berkeley.edu/~necula/ Papers/FloydMeaning.pdf.

[38] [Hoare 69] C A. R. Hoare. An Axiomatic Basis for Computer Programming. Communications Of the ACM, 12(10):576–580, ACM Press, New York (NY, USA), octobre 1969. ISSN 0001-0782. http://dx.doi.org/10.1145/363235.363259.

[39] [Dijkstra 76] Edsger W. Dijkstra. A Discipline of Programming. Series in Automatic Computation. Prentice Hall PTR, Upper Saddle River (NJ, USA), 1976. ISBN 978-0-132-15871-8.121 Bibliographie

[40] [Diaz 82] Michel Diaz. Modelling and Analysis of Communication and Cooperation Protocols Using Petri Net Based Models. In Proceedings of the IFIP WG6.1 Second International Workshop on Protocol Specification, Testing and Verification, pages 465–510, Amsterdam (The Netherlands). North-Holland Publishing Co., 1982. ISBN 0-444-86481-4.

[41] [Davis 88] Alan M. Davis. A Comparison of Techniques for the Specification of External System Behavior. Communications of the ACM, 31(9):1098–1115, ACM Press, New York (NY, USA), septembre 1988. ISSN 0001-0782.

[42] [Harel et al. 90] David Harel, Amir Pnueli, HagiLachover, AmnonNaamad, Michal Politi, Rivi Sherman, AharonShtull-Trauring, ET Mark Trakhtenbrot. STATEMATE: A Working Environment for the Development of Complex Reactive Systems. IEEE Transactions On Software Engineering, 16(4):403–414, IEEE Press, Piscataway (NJ, USA), avril1990.

http://dx.doi.org/10.1109/32.54292.

[43] [Clarke ET Emerson 08] Edmund M. Clarke et E. Allen Emerson. Design and Synthesis of Synchronization Skeletons Using Branching-Time Temporal Logic. In OrnaGrumberg et Helmut Veith, editors: 25 Years of Model Checking, volume 5000 de Lecture Notes in Computer Science, pages 196–215. Springer Berlin Heidelberg, 2008. ISBN 978-3-540-69849-4. http://dl.acm.org/citation.cfm?id=648063.747438.

[44] [Queille et Sifakis 82] Jean-Pierre Queille et Joseph Sifakis. Specification and Verification of Concurrent Systems in CESAR. In MariangiolaDezani-Ciancaglini et Ugo Montanari, éditeurs : Proceedings of the 5th Colloquium on International Symposium on Programming, Volume 137 de Lecture Notes in Computer Science, pages 337–351, London (GB). Springer Berlin Heidelberg, 1982. ISBN 978-3-540-11494-9. http://dl.acm.org/citation. cfm?id=647325.721668.

[45] [Roper 92] Marc Roper. Software Testing: A Selected Annotated Bibliography. Software Testing, Verification and Reliability, 2(3):113–132, John Wiley & Sons, Inc., New York (NY, USA), 1992. ISSN 1099-1689. http://dx.doi.org/10.1002/stvr.4370020303.

[46] [Weyuker 82] Elaine J. Weyuker. On Testing Non-Testable Programs. The Computer Journal, 25(4):465–470, 1982. http://dx.doi.org/10.1093/comjnl/25.4.465.

[47] [Waeselynck 93] Hélène Waeselynck. Vérification de logiciels critiques par le test statistique. Thèse de doctorat, Institut National Polytechnique (INP) de Toulouse, Toulouse (France), 19 janvier 1993. http://homepages.laas.fr/waeselyn/papers/ these_Helene_WAESELYNCK.pdf.130

[48] [Duran et Ntafos 84] Joe W. Duran et Simeon C. Ntafos. An Evaluation of Random Testing. IEEE Transactions on Software Engineering, 10(4):438–444, IEEE Press, Piscataway (NJ, USA), juillet 1984. ISSN 0098-5589. http://dx.doi.org/10.1109/TSE.1984.5010257.

[49] [Thevenod-Fosse 91] Pascale Thevenod-Fosse. Software Validation by Means of Statistical Testing: Retrospect and Future Direction. In AlgirdasAvižienis et Jean-Claude Laprie, éditeurs : Proceeding of the 1st IFIP International Working Conference on Dependable

Computing for Critical Applications (DCCA), volume 4 de Dependable Computing
And Fault-Tolerant Systems, pages 23–50, Santa Barbara (CA, USA). Springer
Vienna (Autriche), 1991. ISBN 978-3-7091-9125-5. http://dx.doi.org/10.1007/978-3-7091-9123-1_2.

[50] [ASSC02] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless Sensor networks: a survey. Comput. Netw. 38(4):393–422, 2002.

[51] [AKK04] J. N. Al-Karaki and A. E. Kamal. Routing techniques in wireless sensor Networks: a survey. In IEEE Wireless Comm., volume 11, pages 6–28, 2004.

[52] [HKB99] Wendi RabinerHeinzelman, Joanna Kulik, and Hari Balakrishnan. Adaptive Protocols for information dissemination in wireless sensor networks.
In MOBICOM, pages 174–185, 1999.

[53] [PPG05] Bryan Parno, Adrian Perrig, and Virgil D. Gligor. Distributed detection of Node replication attacks in sensor networks. In IEEE Symposium on Security
And Privacy, pages 49–63. IEEE Computer Society, 2005.14 Submitted to SAR-SSI 2008
Etat de l'art sur la sécurité dans les réseaux de capteurs sans fil

[54] [WGS+05] Xun Wang, Wenjun Gu, Kurt Schosek, SriramChellappan, and Dong Xuan. Sensor network configuration under physical attacks. In Xicheng Lu andWei Zhao, editors, ICCNMC, volume 3619 of Lecture Notes in Computer Science, pages 23–32. Springer, 2005.

[55] [HBH04] C. Hartung, J. Balasalle, and R. Han. Node compromise in sensor networks: The need for secure systems. Technical Report CU-CS-988-04, Department of Computer Science, University of Colorado at Boulder, 2004.

[56] [WS02] A.D. Wood and J.A. Stankovic. Denial of services in sensor networks. IEEE Computer, October 2002.

[57] [KW03] Chris Karlof and David Wagner. Secure routing in wireless sensor networks: Attacks and countermeasures. Ad Hoc Networks, 1(2-3):293–315, 2003.

[58] [SA99] Frank Stajano and Ross J. Anderson. The resurrecting duckling: Security Issues for ad-hoc wireless networks. In Bruce Christianson, Bruno Crispo,
James A. Malcolm, and Michael Roe, editors, Security Protocols Workshop,
Volume 1796 of Lecture Notes in Computer Science, pages 172–194. Springer, 1999.

[59] [KW03] Chris Karlof and David Wagner. Secure routing in wireless sensor networks: Attacks and countermeasures. Ad Hoc Networks, 1(2-3):293–315, 2003.

[60] [NSSP04] J. Newsome, E. Shi, D. Song, and A. Perrig. The sybil attack in sensor networks: Analysis & defenses. In Information Processing in Sensor Networks, 2004. IPSN 2004. Third International Symposium on, pages 259–268, 2004.

[61] [TCV07] Michel Abdalla Thomas Claveirole, Marcelo Dias De Amorim and Yannis Viniotis. Résistance contre les attaques par capture dans les réseaux decapteurs. In JDIR, 2007.

[62] [DHM05] Jing Deng, Richard Han, and Shivakant Mishra. Countermeasures against Traffic analysis attacks in wireless sensor networks. In SECURECOMM '05: Proceedings of the First International Conference on Security and Privacy For Emerging Areas in Communications Networks, pages 113–126, Washington, DC, USA, 2005. IEEE Computer Society.

[63] [ZSJ03] Sencun Zhu, Sanjeev Setia, and Sushil Jajodia. Leap - efficient security Mechanisms for large-scale distributed sensor networks. In Ian F. Akyildiz, Deborah Estrin, David E. Culler, and Mani B. Srivastava, editors, SenSys, Pages 308–309. ACM, 2003.

[64] [BLM07] ChakibBekara and Maryline Laurent-Maknavicius. A new resilient key Management protocol for wireless sensor networks. In Damien Sauveron, ConstantinosMarkantonakis, Angelos Bilas, and Jean-Jacques Quisquater, Editors, WISTP, volume 4462 of Lecture Notes in Computer Science, pages 14–26. Springer, 2007.

[65] [GSJ+03] Marco Gruteser, Graham Schelle, Ashish Jain, Richard Han, and Dirk Grunwald. Privacy-aware location sensor networks. In Michael B. Jones, editor, HotOS, pages 163–168. USENIX, 2003.Submitted to SAR-SSI 2008 13 David MARTINS et Hervé GUYENNET

[66] [LND05] Donggang Liu, Peng Ning, and Wenliang Du. Detecting malicious beacon Nodes for secure location discovery in wireless sensor networks. In ICDCS,

Pages 609–619. IEEE Computer Society, 2005.

[67] [YZV03] Z. Yan, P. Zhang, and T. Virtanen. Trust evaluation based security solution In ad hoc networks. In NordSec 2003, Proceedings of the Seventh Nordic Workshop on Secure ITS Systems, 2003.

[68] [ZBDK04] H. Zhu, F. Bao, R. H. Deng, and K. Kim. Computing of trust in wireless Networks. In Proceedings of 60th IEEE Vehicular Technology Conference, Los Angles, California, September 2004.

[69] [NV07] Pissinou Niki and Crosby Garth V. Cluster-based reputation and trust for Wireless sensor networks. In Consumer Communications and Networking Conference, 2007. CCNC 2007. 4th IEEE, pages 604–608, January 2007.

[70] [RLW+04] Kui Ren, Tieyan Li, Zhiguo Wan, Feng Bao, Robert H. Deng, and Kwangjo Kim. Highly reliable trust establishment scheme in ad hoc networks. Computer Networks, 45(6):687–699, 2004.

[71] [GS04] Saurabh Ganeriwal and Mani B. Srivastava. Reputation-based framework For high integrity sensor networks. In Sanjeev Setia and VipinSwarup, Editors, SASN, pages 66–77. ACM, 2004.

[72] [OZ07] Vladimir Oleshchuk and Vladimir Zadorozhny. Trust-aware query processing in data intensive sensor networks. In SENSORCOMM ’07: Proceedings of The 2007 International Conference on Sensor Technologies and Applications, Pages 176–180, Washington, DC, USA, 2007. IEEE Computer Society.

[73] [LS05] Zhengqiang Liang and Weisong Shi. Pet: A personalized trust model with Reputation and risk evaluation for p2p resource sharing. In HICSS. IEEE Computer Society, 2005.

[74] [RR06] Lopez J. Roman R., Jianying Zhou. Applying intrusion detection systems To wireless sensor networks. In Consumer Communications and Networking Conference, 2006. CCNC 2006. 3rd IEEE, pages 640–644, January 2006.

[75] L’Internet des objets a besoinde plates-formes sécurisées DaveKleidermacher,directeur technique De Green Hills Software.L’EMBARQUÉ / N°7 / 2014

[76] H. Janzadeh, K. Fayazbakhsh, M. Dehghan, M.S. Fallah, “A securecredit- based Cooperations timulating mechanism for MANETsusing hash chains”. Future Generation Computer Systems, Volume 25 Issue 8, September, 2009. [1] http://www.internet-of-things-research.eu/ (consulté en Novembre 2015).

[77] [John R. Douceur 02]. The sybil attack. In Peer-to-Peer Systems, First Inter-national Workshop, IPTPS 2002, Cambridge, MA, USA, March 7-8, 2002, Revised Papers, pages 251{260. Springer, 2002.

[78] Brian Neil Levine, Clay Shields, and N. Boris Margolin. A survey of solutionsto the sybil attack. Tech report 2006-052, University of Massachusetts Amherst, Amherst, MA, October 2006.

[79] James Newsome, Elaine Shi, Dawn Xiaodong Song, and Adrian Perrig. The sybil attack in sensor Networks: analysis & defenses. In Proceedings of the Third International Symposium on Information Processing in Sensor Networks, IPSN 2004, Berkeley, California, USA, April 26-27, 2004, pages 259{268. ACM, 2004.

[web1] "Big Data Analytics", http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics, Mars 2017

[web2] Paul Peton, "Les 4 sources du BigData", http://www.communication-web.net/2016/03/07/les-4-sources-du-big-data/, 7 mars 2016

[web3] "Qui sont les acteurs du marché Big Data?", https://www.mba-esg.com/actus/acteurs-big-data, 2014

[web4] Dean, Jeffrey &Ghemawat, Sanjay, "Mapreduce" http://dictionnaire.sensagent.leparisien.fr/MapReduce/fr-fr, 2004

[web5] Laurent, A. (2011, juillet 13). "VMWare lance vSphere 5 et sa nouvelle suite dédiée à la gestion du cloud. " Consulté le novembre 19, 2011, sur Clubic: http://pro.clubic.com/it-business/actualite-435346-vmware-vsphere-5-dediee-gestion-cloud.html

[web6] Elyan, J. (2010, juillet 07). "Ne confondez pas Cloud et SOA explique Gartner. Consulté le Décembre 2011, sur Le monde du cloud:"http://www.lemondeducloud.fr/lire-ne-confondez-pas-soa-et-cloud-explique-le-gartner-31308.html

[web7] Oeillet, A. (2011, septembre 08). "Google : l'email dans le cloud a des bienfaits énergétiques. Consulté le septembre 08, 2011, sur Clubic: "http://www.clubic.com/internet/google/actualite-445486-google-email-cloud-bienfaits-energetiques.html

[Web 8]"Web 1.0 – Web 2.0 – Web 3.0" Sepetember 2008 http://www.atelier-informatique.org/internet/evolution-web-10-web-20-web-30/358/

[Web9] https://www.ciscomadesimple.be/2014/11/08/protocole-ipv4-les-bases/

[Web10] https://aruco.com/2014/08/infographie-internet-objets/

[Web11] http://www.smartgrids-cre.fr/index.php?p=objets-connectes-definition

[Web12] http://www.journaldunet.com/ebusiness/expert/58395/tirez-profit-de-l-internet-des-objets-grace-Aux-plateformes-irm.shtml

[Web13] http://www.usine-digitale.fr/editorial/l-internet-des-objets-un-marche-de-seulement-550-milliards-De-dollars-en-2025.N394582

[Web14] http://www4.ac-nancy-metz.fr/Etablissement/MOSELLE/College/clg-mermoz-yutz/technologie/Cours/S11-c/S11-c-ressources.pdf

[Web15] https://www.france-science.org/Les-Objets-Connectes-la-nouvelle.html

[Web16] http://www.atelier.net/trends/articles/internet-objets-besoin-de-normes-se-diffuser_425248

[Web17] "Quelques exemples intéressants de ce que fait l'IoT dans le domaine de la santé et du transport" 2015 http://www.objetconnecte.com/internet-objets-transports-sante-1109/

[Web18] "Cisco and/or its affiliates "2013 https://iot.goffinet.org/iot_modeles_et_architectures.html

[Web19] http://www.leparisien.fr/flash-actualite-economie/internet-des-objets-la-bataille-des-futurs-standards-de-communication-est-engagee-17-07-2015-4952159.php

[Web20] http://www.automation-sense.com/blog/blog-objets-connectes/reseaux-protocoles-iot-m2m.html

[Web21] www.ordinateur.cc/réseaux/Autre-Réseaux-informatiques/78561.html

[Web22] http://www.ordinateur.cc/r%C3%A9seaux/Autre-R%C3%A9seaux-informatiques/78561.html

[Web23] http://yannickh.chez.com/protocoles.html

[Web24] https://www.ciscomadesimple.be/2014/11/08/protocole-ipv4-les-bases/

[Web25] http://www.commentcamarche.net/contents/524-le-protocole-ipv6

[Web26] http://www.techniques-ingenieur.fr/base-documentaire/technologies-de-l-information-th9/internet-des-objets-42612210/6lowpan-te8002/protocole-6lowpan-te8002niv10003.html

[Web27] http://www.commentcamarche.net/faq/7294-tcp-udp-quelles-differences

[Web28] http://yannickh.chez.com/protocoles.html

[Web29] http://www.materiel-informatique.be/upnp.php

[Web30] https://www.arrow.com/fr-fr/research-and-events/articles/protocols-for-the-internet-of-things

[Web31] https://fr.wikibooks.org/wiki/D%C3%A9butez_dans_XMPP/Qu%27est-ce_qu%27XMPP_%3F

[Web32] http://blog.octo.com/modeles-architectures-internet-des-objets/

[web33] http://www.commentcamarche.net/contents/47-piratage-et-attaques-informatiques

[web34] http://www.gemalto.com/france/iot/securite-iot

[Web35] http://www.oezratty.net/wordpress/2016/peut-on-securiser-internet-objets/

[Web36] http://bfmbusiness.bfmtv.com/entreprise/combiner-securite-hard-et-soft-un-nouveau-defi-pour-l-Internet-des-objets-1051946.html

[Web37] http://www.iot-butler.eu/ (consulté en Novembre 2015).

[Web 38] http://www.hydramiddleware.eu/news.php (consulté en Novembre 2015).

[Web 39] https://www.nitrd.gov/ (consulté en Novembre 2015).

[Web 40] http://www.internet-of-things-research.eu/ (consulté en Novembre 2015).

[Web 41] www.theses.fr/2015VERS012V.pdf

[Web 42] https://tel.archives-ouvertes.fr/tel-01166047/document

[Web 43] https://aruco.com › Articles

[Web 44] https://tel.archives-ouvertes.fr/tel-01143685/document

[Web 45] https://tel.archives-ouvertes.fr/tel-00458244v1/document