

Développement d'un pare-feu domestique

Pré-rapport de projet

Activité d' Apprentissage S-INFO-037

Rémy Decocq

Année Académique 2018-2019
Master en Sciences Informatiques, bloc 1
Faculté des Sciences, Université de Mons

Table des matières

1 Présentation de l'<i>Internet des Objets</i>	5
1.1 Généralités	5
1.2 Caractéristiques des équipements de l' <i>IoT</i>	5
1.2.1 Domaines d'application	5
1.2.2 L'environnement <i>smarthome</i>	6
1.2.3 Restrictions des équipements	7
1.3 Protocoles adaptés à l' <i>IoT</i>	7
2 Présentations des pare-feux	9
2.1 Généralités	9
2.2 Différentes positions et fonctions dans l'architecture réseau	9
2.2.1 Les pare-feux niveau réseau	9
2.2.2 Les pare-feux niveau hôte	10
2.3 Les types de pare-feu	11
2.3.1 Pare-feu de filtrage (sans état)	11
2.3.2 Pare-feu à état	12
2.3.3 Pare-feu applicatif	12
2.3.4 Pare-feu applicatif proxy	13
2.3.5 Pare-feu identifiant	13
2.4 Systèmes de détection d'intrusion (<i>IDS</i>)	14
2.4.1 Les NIDSs	14
2.4.2 Les HIDSs	15
2.4.3 Les DIDSs	15
2.5 Systèmes de prévention d'intrusion (<i>IPS</i>)	16
2.6 Les pare-feux nouvelle génération	16
2.7 Les pare-feux de référence	17
2.7.1 Sous forme de software	17
2.7.2 Sous forme d'appliance	18
3 La sécurité dans l'<i>IoT</i>	20
3.1 Motivations et exemples	20
3.2 Description des besoins en terme de sécurité	21
3.3 Vulnérabilités liées aux équipements IoT	22
3.4 Architecture globale des systèmes impliquant des équipements IoT	23
3.5 Les archétypes d'attaques courantes	25
3.5.1 Sur la couche perception	25
3.5.2 Sur la couche réseau	25
3.5.3 Sur la couche support et application	26
3.6 Ressources, solutions et outils existants	27
3.6.1 Le travail d'OWASP	27
3.6.2 La standardisation de protocoles adaptés sécurisés	27
3.6.3 La détection des équipements en réseau et de leurs vulnérabilités	27
4 Développement d'une application <i>IoTMonitor</i> - introduction et concepts	30
4.1 Présentation générale	30
4.1.1 Contexte et environnement ciblé	30
4.1.2 Objectifs de l'application	30
4.2 Composants du système de surveillance	32
4.2.1 Modules actifs et passifs - abstraction des programmes	32
4.2.2 Routine - automatisation de la surveillance	33
4.2.3 Netmap - recensement des équipements du réseau	34
4.2.4 Centre d'événements - logging et lancement d'alertes	35
4.3 Modèle d'abstraction des logiciels	35
4.3.1 Modèle général d'abstraction	35

4.3.2	Design pour un module actif	36
4.3.3	Design pour un module passif	37
5	Implémentation de l'application	38
5.1	Justification des choix généraux	38
5.2	Discussions des choix d'implémentation	38
5.2.1	Gestion du temps	38
5.2.2	Exécution des logiciels	38
5.2.3	Parsing des sorties	38
5.2.4	Stockage et agrégation des informations	38
5.2.5	Interactivité avec l'application	38
5.3	Sélection des logiciels à intégrer	38
6	Manuel d'utilisation	39
7	Tests et évaluation de l'application	40
7.1	Simulation d'équipements	40
7.2	Déploiement dans un milieu domestique	40
7.3	Stress-test dans un environnement peuplé	40

Introduction

Depuis maintenant plusieurs années, la connectivité n'a cessé d'évoluer : en se limitant au domaine de l'Internet entre 2000 et 2015, une estimation de l'augmentation du pourcentage de la population mondiale l'utilisant avoisine 40% [6] [3]. Que ce soit dans le cadre d'infrastructures de type « mainframe » ou dans le contexte des ordinateurs personnels, les technologies et équipements relatifs au réseau et aux communications sont devenus indispensables. En raison du potentiel croissant d'interconnexion, les ordinateurs et équipements mis en réseau s'exposent à davantage de menaces. Heureusement, parallèlement à cette évolution, les performances de ces machines rejoignant de nouveaux réseaux ont également suivi une amélioration en terme de performances. Cela a permis d'en renforcer la sécurité à plusieurs niveaux et surtout d'intercepter efficacement les menaces étrangères liées à l'utilisation des réseaux. À l'heure actuelle, les OS utilisés classiquement sur des machines desktop fournissent un pare-feu simple (*Windows Defender*, un utilitaire fourni de base dans MAC OS X, *iptables/Netfilter* ou autre pour les distributions Linux). Ce dernier tourne en arrière plan de façon quasi invisible car il demande peu de ressources par rapport à ce qu'une machine actuelle peut offrir.

En parallèle avec la montée en puissance de ces machines de type desktop, serveurs, etc. s'est développée depuis à peu près les années 2000 la tendance de l'« Internet des Objets », ou encore plus communément abrégé IoT pour *Internet of Things*. Bien qu'assez large, cette dénomination regroupe beaucoup d'objets et de concepts, qu'ils soient virtuels ou non mais possédant un dénominateur commun : la capacité de communiquer en réseau avec d'autres équipements. Cela englobe par exemple la domotique, les outils et capteurs de mesures diverses, les imprimantes et scanneurs en réseau, etc. Tous ces éléments convergeraient idéalement vers une mise en réseau commune, leur permettant de communiquer malgré leur nature hétérogène [17]. Cela peut se faire via le réseau Internet, il n'est pas rare d'orienter ces connections vers un cloud permettant de traiter globalement et intelligemment la masse de données qu'il reçoit de ces équipements [21]. Or, comme évoqué ci-dessus, plus on s'interconnecte et plus on s'ouvre à des attaquants potentiels, ce qui pose problème si rien n'est mis en place pour s'en protéger.

Ce travail aura pour objectif premièrement de faire un état de l'art des dispositifs de protection qui sont actuellement déployés dans l'IoT et plus particulièrement dans le cadre domestique. Il s'agit d'un monde beaucoup plus hétérogène et restreint en terme de ressources que celui des ordinateurs qu'on retrouve classiquement dans ce milieu, de fait il n'est pas toujours possible de réutiliser telles quelles toutes les technologies de protection y attenant. De plus, il faut prendre en considération les spécificités d'une habitation : au centre de tous ces équipements communiquant se trouve l'habitant, une personne n'étant pas forcément apte à manipuler et contrôler ces nouvelles technologies. Il sera également question d'établir une vision globale des différentes dispositifs de pare-feux existant et sous quelles formes ceux-ci peuvent être implantés dans les équipements d'une habitation classique. Deuxièmement, il sera question de mettre en pratique ces connaissances pour développer un système en lien avec ces nouvelles mesures de sécurité inhérentes à l'IoT. Celui-ci fera intervenir les connaissances acquises au préalable sur les pare-feux.

1 Présentation de l'*Internet des Objets*

1.1 Généralités

L'Internet des objets, qu'on désignera par *IoT* pour le terme plus répandu de *Internet of Things*, est une notion englobant un vaste ensemble de dispositifs. Aucune définition formelle n'est acceptée globalement, mais plusieurs organismes ont tenté d'en établir une ébauche. Par exemple, l'ITU-T (ITU Telecommunication Standardization Sector) Y.2060 le définit comme tel :

« *Global infrastructure for the society, enabling advanced services by inter-connecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.* »

Derrière cette définition très générale, on peut distinguer plusieurs sous-groupes d'objets connectés distincts, aux applications tant variées que hétérogènes, dans des domaines et secteurs également très différents. C'est ce qui fait la force et en même temps la faiblesse de cet ensemble d'équipements et services qu'on regroupe derrière le terme *IoT* et que des efforts considérables sont déployés pour inter-connecter au maximum. C'est un domaine d'étude intéressant car il représente littéralement ce qu'on pourrait considérer comme le futur de notre environnement technologique. De fait, les équipements que l'on peut associer à une partie de l'*IoT* ont déjà fait leur apparition dans notre quotidien : en 2017 on comptait 8,4 milliards de machines en présentant les caractéristiques et les estimations pour l'année 2020 tendent vers 20,4 milliards d'objets connectés d'après la société d'analyse Gartner [12].

1.2 Caractéristiques des équipements de l'*IoT*

1.2.1 Domaines d'application

Les secteurs dans lesquels l'*IoT* s'est implanté ces dernières années sont nombreux et très variés : ils s'étendent de l'industrie au domaine des soins de santé en passant par la tendance des *Smart Home*. C'est ce dernier domaine qui est approfondi dans ce travail et qui sera le plus sous-entendu par la suite quand le terme *IoT* est utilisé. La Figure 1 présente une vision d'un schéma global des autres domaines qui gravitent autour du vaste monde de l'*IoT*.

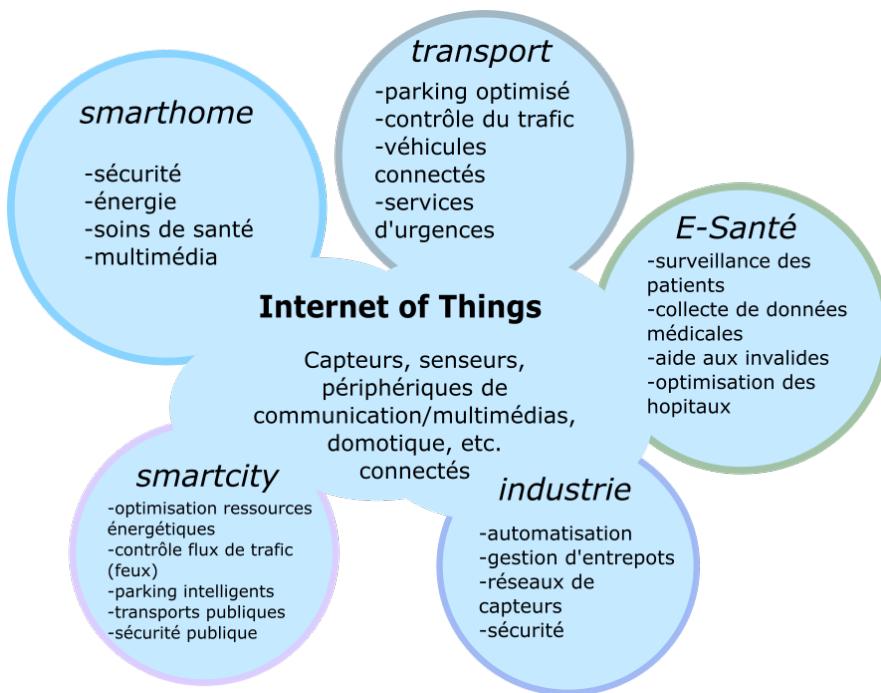


FIGURE 1 – Domaines d'application de l'*IoT*

1.2.2 L'environnement smarthome

Le terme émergeant *Smart Home/smarthome* est une fois de plus très englobant et général. Il n'en existe pas de définition formelle et communément acceptée. Basman M. Hasan et al. [10] en présentent plusieurs. Un résultat les unifiant pourrait être

« Une smarthome est un environnement lié au domicile particulier où plusieurs équipements ou sous-systèmes sont inter-connectés et où les informations qu'ils échangent sont collectées et utilisées afin de surveiller, réguler et automatiser l'écosystème du domicile ».

L'utilisateur en tant que personne physique y vivant est donc au centre de cette architecture, et y siège comme le principal intervenant. Effectivement, l'objectif global du déploiement de tous ces dispositifs est l'amélioration de sa qualité de vie. La notion d'intelligence est intrinsèquement liée avec celle de l'interconnexion de tous ces capteurs et actuateurs déployés dans l'environnement du domicile. Il s'agit d'en récolter et regrouper toutes les données en un point central doté d'une capacité de traitement plus évoluée afin qu'il puisse en tirer une optimisation globale de l'habitation (du système de sécurité, des économies d'énergie, de temps par l'automatisation, etc.).

Une certaine classification fonctionnelle peut être établie pour distinguer de façon plus concrète les différents équipements qui peuvent intervenir dans l'écosystème d'une smarthome. Elle est schématisée par la Figure 2, inspirée de [10]. Ce qui est désigné par *point d'interconnexion* peut dépendre de l'architecture réelle d'une smarthome [21]. Dans la plupart des cas il s'agit d'une machine faisant office de collecteur pour toutes les données transitant dans le domicile et de gateway vers le reste de l'Internet, éventuellement le cloud associé.

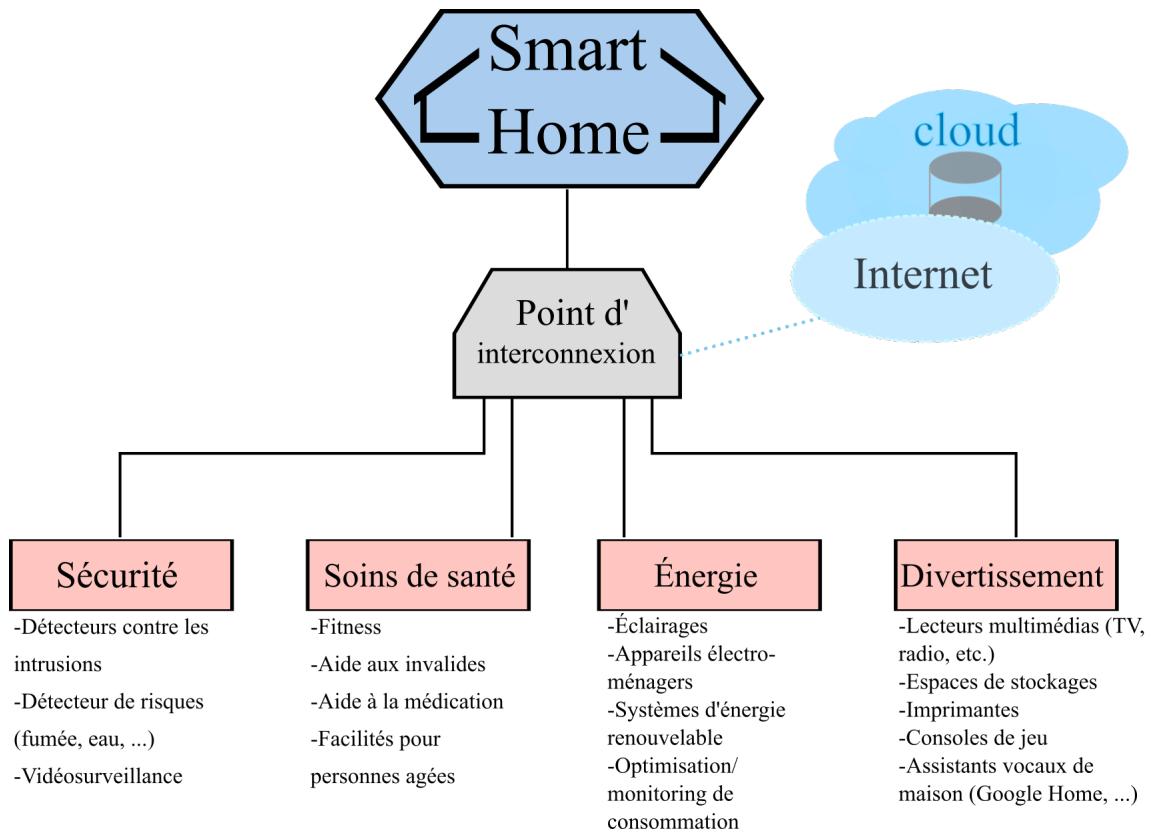


FIGURE 2 – Classification fonctionnelle des équipements IoT d'une smarthome

1.2.3 Restrictions des équipements

Malgré le fait que l'ensemble des objets considérés comme appartenant à l'IoT soit très hétérogène, on peut distinguer plusieurs caractéristiques communes à beaucoup d'entre eux. Elles tendent généralement vers ce qui est vu comme une restriction par rapport à un ordinateur type classique (*desktop*). Ces éléments constituent les plus gros freins au développement de la sécurité sur de tels système [9]. Les conséquences de ces restrictions sont discutées plus en détail dans la section 3 de ce document.

Conçus pour satisfaire une unique fonction

Le meilleur exemple est celui des capteurs. Un capteur a pour objectif de faire une mesure d'une grandeur physique (température, pression, etc.), d'en tirer une valeur numérique et de faire remonter via son interface avec le réseau cette information vers une unité centrale. Cette dernière accumule ainsi des mesures provenant des nœuds distribués pour y appliquer un traitement, et c'est à ce plus haut niveau que le processus de décision a lieu s'il est nécessaire. Ce genre d'équipement est généralement minimaliste au possible et ne peut donc pas remplir d'autre tâche.

Requièrent une faible consommation énergétique

Les systèmes embarqués n'ont pas toujours accès à une source illimitée d'énergie, et auront donc une durée de vie limitée à celle de leur batterie. En conséquence, il est souhaitable d'économiser un maximum, ce qui peut se faire en réduisant les temps d'éveil de l'équipement et en optimisant le nombre d'opérations effectuées quand il tourne à plein régime. Dès lors, certains protocoles et algorithmes doivent être adaptés (relatifs aux communications réseaux mais aussi à la sécurité) [21].

Sont contraints en ressources CPU, mémoire et radio

Ces contraintes sont aussi identiques à celles des systèmes embarqués classiques. En plus de celles-ci, on peut mettre en évidence le fait que les radios (quand il s'agit d'une interface sans fil) sont assez faibles et ne permettent pas des communications à haut débit. En résulte également que des techniques comme le saut de fréquence et les algorithmes de chiffrement de type asymétrique sont plus compliquées à mettre en œuvre [19].

Sont programmés à un bas niveau d'abstraction

Étant conçus pour ne remplir que des fonctions spécifiques, certains équipements ne sont pas programmables en utilisant des langages de haut niveau. Par conséquent, ce sont souvent des boîtes noires difficilement manipulables et statiques : les mises à jour et patch de sécurité ne sont pas déployables aisément par les constructeurs [21]. Les seules interactions que l'utilisateur lambda peut avoir avec l'équipement sont celles prévues par l'interface de ce dernier s'il y en a une (pour le configurer, consulter son état, etc.).

1.3 Protocoles adaptés à l'IoT

Les restrictions communément présentées par les équipements IoT énumérées dans la sous-section 1.2.3 imposent de repenser les protocoles utilisés dans un réseau non soumis à ces contraintes. Les protocoles décrits ci-dessous ont été conçus dans cette optique et sont utilisables dans un réseau composé d'équipements restreints en ressources, tel qu'un réseau de capteurs communiquant via la technologie sans-fil (WSN - *Wireless Sensors Network*). La figure 3 en présente un récapitulatif, mettant en exergue les différences avec les protocoles utilisés dans un réseau classique (supportant la pile *TCP/IP*).

Le standard **IEEE 802.15.4** a été défini par l'IEEE (Institute of Electrical and Electronics Engineers) pour les réseaux sans-fil dits « personnels » (WPANs - *Wireless Personal Area Networks*) c-à-d destinés à couvrir une zone d'émission de l'ordre de quelques dizaines de mètres. IEEE 802.15.4 supporte les couches physiques (PHY) et lien (MAC - *Medium Access Control*) dans ce cadre, en considérant le manque de ressources des équipements [14] [21].

Les réseaux utilisant IEEE 802.15.4 dans les couches plus basses sont contraints par la taille maximum que les trames peuvent contenir (pour les couches supérieures) selon ce standard, c-à-d 102 bytes. Sachant que la taille maximum d'un paquet IPv6 peut aller jusqu'à 1280 bytes par défaut, pour utiliser ce protocole il est nécessaire d'introduire un mécanisme d'adaptation. C'est l'objectif principal de la couche intermédiaire **6LoWPAN** : elle gère la fragmentation et le ré-assemblage des paquets IPv6 pour qu'ils puissent être transférés dans un réseau utilisant IEEE 802.15.4.

6LoWPAN propose d'autres services tels que la compression des en-têtes IPv6 (intuitivement nécessaire) et de l'auto-configuration de l'équipement joignant un réseau (similaire à IPv6).

L'aspect routage (sous-entendu pour des adresses IPv6) est généralement géré par le biais de **RPL** (bien qu'il soit possible de le faire avec 6LoWPAN). Il s'agit d'un protocole à vecteurs de distances adapté aux réseaux comme les WSN, comme le suggère son nom complet : *Routing Protocol for Low-Power and Lossy Networks*. Il se base sur une topologie virtuelle en arbre établie par un DODAG (*Destination Oriented Directed Acyclic Graph*), la racine étant un nœud qui a pour tâche de collecter toutes les informations montantes (et de les transmettre à un nœud extérieur au WSN). Plusieurs métriques peuvent être envisagées pour déterminer les routes, comme le nombre de sauts, la qualité des sources d'énergie de chaque nœud, les probabilités de réussite de transmission entre les nœuds. RPL est donc assez souple que pour s'adapter aux besoins et aux particularités des réseaux de natures variées que l'on peut retrouver dans l'IoT.

DTLS (*Datagram Transport Layer Security*) est le mécanisme de sécurisation sur lequel peuvent se reposer les applications utilisées par les équipements IoT. Son homologue dans les réseaux non restreints est TLS, qui se repose sur la fiabilité de transport garantie par TCP : si un paquet est pas ou mal délivré, la connexion est interrompue [18]. DTLS est conçu pour fonctionner par dessus **UDP**, qui est le protocole *de facto* utilisé pour le transport dans les réseaux IoT (celui-ci demandant moins de ressources). Il n'est pas forcément destiné à l'IoT, aussi des implémentations ont vu le jour dans cette optique ([TinyDTLS](#)).

Du côté applicatif, **CoAP** est un bon candidat. Ce protocole est maintenu par l'IETF, par l'équipe du CoRE (*Constrained RESTful Environments*) et est de fait basé sur l'architecture REST [14] comme son homologue HTTP. Il s'agit donc d'un modèle client-serveur. CoAP aspire à restreindre le dialecte HTTP à un sous-ensemble adapté aux contraintes des équipements IoT (aux communications 6LoWPAN) et est prévu pour fonctionner sur UDP. Ce dernier n'étant pas un protocole de transport fiable, CoAP implémente un mécanisme simple impliquant des messages *Confirmables* et *Acknowledge* pour pouvoir assurer un minimum de fiabilité si les données à transmettre sont critiques.

MQTT est un autre protocole qui diffère de CoAP sur plusieurs points¹. C'est un protocole de messagerie orienté souscription/publication léger par rapport à HTTP, la taille des messages se restreignant à quelques bytes contre plusieurs centaines [20]. L'architecture est également différente : on a un équipement dit *broker* qui joue le rôle de serveur de messagerie autour duquel s'articulent plusieurs clients désirant communiquer entre eux. Chaque client s'abonne à un *topic* pour lequel il est intéressé en établissant une liaison avec le broker, et recevra un message concernant ce topic posté par tout autre nœud. Cette liaison, ouverte en permanence, utilise TCP et peut donc être authentifiée et chiffrée avec TLS. En plus du fait que la longueur des noms des topics peut poser problème avec la taille maximum des frames 802.15.4, cette utilisation de TCP pose problème pour les nœuds trop limités en ressources. Un port vers UDP du nom de **MQTT-SN** a été établi à cette fin.

Couches du modèle	<i>Protocoles de la stack IoT</i>	<i>Protocoles de la stack TCP/IP</i>
Application	IETF CoAP MQTT ...	HTTP FTP DNS ...
Sécurité	DTLS	TLS
Transport	UDP	TCP UDP
Réseau	IPv6 IETF RPL	IPv4 IPv6
Adaptation	IETF 6LoWPAN	Non défini
Lien	IEEE 802.15.4 MAC	Dépend du média et de la technologie (IEEE 802.11 WiFi, Ethernet, ...)
Physique	IEEE 802.15.4 PHY	

FIGURE 3 – Comparaison du modèle en couche classique (TCP/IP) avec celui adapté à l'IoT

1. https://www.eclipse.org/community/eclipse_newsletter/2014/february/article2.php

2 Présentations des pare-feux

2.1 Généralités

Un pare-feu est un dispositif, virtuel ou matériel, qui surveille et contrôle le lien entre un réseau dit de confiance et un réseau extérieur non fiable. Typiquement, ce réseau potentiellement dangereux est Internet et la zone à protéger est le réseau interne d'une entreprise ou d'une habitation. L'existence des pare-feux est une conséquence du besoin d'outils de protection aux bordures de réseaux distincts contrôlés par des entités différentes. D'une part, il faut garantir que des données internes confidentielles restent à l'intérieur du réseau de confiance. D'autre part, il est nécessaire de filtrer les données entrantes et sortantes de ce réseau de sorte qu'aucune menace ne s'y infiltre par des flux, même initiés depuis l'intérieur du réseau de confiance. Ces filtres sont créés et assemblés à partir de *politiques* ou *règles* définies par défaut ou par les personnes compétentes liées à l'entité gérant le réseau.

2.2 Différentes positions et fonctions dans l'architecture réseau

Deux classes de pare-feux sont distinguables en fonction de ce qu'ils visent à protéger. La première correspond aux pare-feux au niveau réseau (*network firewalls*), situés en bordure de LANs, WANs et intranets. Ceux-ci, de par leur nature de barrière entre réseaux, peuvent également fournir des services plus évolués : système de NAT, gestion de *zones démilitarisées* (définies ci-dessous), service DHCP, etc. [15] La seconde classe agit au niveau des noeuds du réseau eux-mêmes (*host-based firewalls*), protégeant une machine physique et non pas un réseau entier. Ces pare-feux se présentent donc sous forme de logiciels, directement intégrés au niveau du système d'exploitation ou installés à un plus haut niveau.

La figure figure 6 schématise la place de ces dispositifs, mettant en scène le scénario simplifié d'un utilisateur surfant sur le site d'une entreprise. Une requête HTTP pour une page est créée et envoyée depuis le domicile, destinée aux serveurs de l'entreprise. Ceux-ci sont protégés derrière une machine dédiée uniquement à la fonction de pare-feu (qu'on qualifie alors par le terme *appliance*), dans une zone tampon désignée par **DMZ** pour *Demilitarized Zone* (zone démilitarisée). Cette zone intermédiaire est un sous-réseau contenant les machines étant destinées à être accédées depuis Internet mais qui n'ont aucun besoin d'accéder au réseau local de l'entreprise (typiquement des serveurs). Ainsi, si une attaque contre ces machines aboutit et que l'attaquant en prend le contrôle, il n'a pas encore un accès direct aux autres machines de l'entreprise (si le pare-feu est toujours fonctionnel). Cela permet en plus d'établir un filtrage plus fin effectué par le pare-feu, qui utilise des règles différentes en fonction de la zone ciblée et donc des services à considérer comme légitimes.

2.2.1 Les pare-feux niveau réseau

Afin de remplir leur fonction, ces pare-feux sont placés en bordure du réseau à protéger et sont directement liés aux machines qui font office de *gateway* (routeurs). Le trafic échangé entre les réseaux ainsi connectés passe donc par le pare-feu afin d'être analysé et filtré, ce qui peut représenter beaucoup de données à traiter. Le pare-feu doit offrir une vitesse de traitement proportionnelle aux débits et à la qualité des liens qui le traversent afin de ne pas être un goulot d'étranglement. C'est pourquoi ces pare-feux doivent être très efficaces et sont communément portés (partiellement) au niveau matériel. On parle de *hardware-based firewall appliances*, qui sont des machines physiques dont le seul objectif est de remplir les tâches d'un pare-feu le plus efficacement possible. On y retrouve deux composants : la partie applicative software ou firmware remplissant la fonction de pare-feu, qui repose sur la deuxième partie plus basse constituée de juste ce qu'il faut d'un OS particulier (*jeOS - just enough Operating System*). Cet OS est généralement propriétaire, lié au fabricant du hardware sur lequel les deux parties opèrent et donc optimisé pour garantir les performances requises.

À plus petite échelle, dans un réseau domestique par exemple, on retrouve également des pare-feux directement implantés dans le routeur qui fait office de *gateway* pour l'habitation. Ceux-ci sont fatalement moins efficaces et complets que les matériels spécialement dédiés à cette unique fonction.

À titre indicatif, la Figure 4 donne un exemple d'une *appliance* pare-feu issue de la gamme des Cisco ASA 55xx-X, la dernière génération que Cisco a introduit sur le marché en 2018 [4]. Il ne s'agit pas du modèle le plus performant de la gamme en comparant les [performances et fonctionnalités](#), mais convient typiquement bien pour être déployé dans une petite entreprise (son prix avoisine 500€). Dans le cas d'une habitation, l'analogue correspondant est dans la majorité des cas le modem distribué par le fournisseur d'accès internet au domicile. La société Proximus fournit par exemple des *b-box*, que leurs clients peuvent configurer par une interface web. Dans ces configurations, il existe bien un onglet intitulé *firewall*, mais celui-ci ne présente que peu d'options de configuration comme le montre la Figure 5. Dans d'autres onglets, on peut trouver d'autres fonctionnalités qui peuvent y être relatives : restriction d'accès pour certains noeuds en fonction de l'heure (par adresse MAC) et blocage de certains sites par domaine. Cela semble assez faible, les utilisateurs du réseau ont tout intérêt à installer sur leurs équipements personnels des pare-feux pour hôtes.

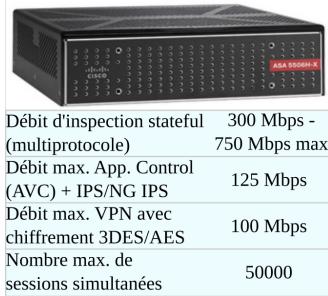


FIGURE 4 – Cisco ASA 5506H-X

Level	Low	Medium	High	?
Low: All connection attempts coming from the WAN or initiated from the LAN are permitted				
Medium:	All connection attempts coming from the WAN are rejected with the exception of those that have been authorized through port forwarding, DMZ or remote access configuration.			
All Services initiated from the LAN are permitted.				
High:	All connection attempts coming from the WAN are rejected with the exception of those that have been authorized through port forwarding, DMZ or remote access configuration.			
Services initiated from the LAN are restricted to the ones authorized via the rules set in the firewall (common services are covered by default).				

FIGURE 5 – Options de config. du pare-feu pour une b-box 3V

2.2.2 Les pare-feux niveau hôte

Du fait que les pare-feux de cette classe sont généralement déployés sur des machines de type desktop, la dénomination *pare-feu personnel* est aussi utilisée. Ces ordinateurs sont généralement munis de tels pare-feux par défaut ou peuvent tout du moins supporter leur installation par l'utilisateur si l'OS utilisé est classique (Windows, MAC OS X, ...). Le pare-feu agit jusqu'au niveau applicatif [16] et sous forme d'un *service* ou d'un *daemon* en fonction du système d'exploitation utilisé. Cela permet une proximité étroite avec l'OS et les autres processus. Le pare-feu personnel est donc capable de contrôler le trafic réseau demandé par chaque application et de détecter les menaces engendrées par certains flux. Ces dernières peuvent être entrantes ou sortantes : une machine extérieure qui tente d'établir une connexion suspecte ou un exécutable sur la machine à protéger qui initie un trafic avec une cible blacklisted, par exemple.

Les principales fonctionnalités d'un pare-feu personnel devraient être au minimum les suivantes [15]

- Bloquer les attaques et comportements dangereux du réseau extérieur : scanning des ports ouverts, attaques par fragmentation, *IP Spoofing*, etc.
- Stopper les menaces venant de l'intérieur : un exécutable comme un *malware* ou un *spyware* qui tente d'établir une connexion vers l'extérieur doit être bloqué et mis en quarantaine
- Présenter de l'automatisation car d'une part car un utilisateur non-expérimenté pour sa configuration doit quand même rester protégé et d'autre part il doit pouvoir se mettre à jour automatiquement
- Agir au niveau applicatif : un malware pourrait utiliser le port web 80 pour se répandre par exemple, pour détecter cela il faut analyser les payloads des paquets et disposer d'une base de données à jour pour y déceler des patterns malicieux sur base de leur signature
- Alerter l'utilisateur quand un événement survient, et le logger avec suffisamment d'informations pour qu'il puisse prendre une décision adaptée
- Éviter les *faux-positifs* (bloquer du trafic légitime)

Que ce soit en entreprise ou dans une habitation, il est donc probable que l'on retrouve des pare-feux appartenant à ces deux classes distinctes là où ils sont efficaces (voir Figure 6). Il ne s'agit que d'échelles différentes, qui impliquent également des intervenants différents. En entreprise, l'administrateur sécurité/réseau configure du matériel spécifique afin de sécuriser les frontières de son réseau tout en préservant ses performances. Il est aussi possible qu'il installe dans les machines internes de son réseau des pare-feux assez puissants pour empêcher les propagations des attaques déclenchées. Dans le domicile, l'habitant est sommairement protégé par le routeur fourni par son FAI qui inclut un pare-feu réseau par défaut. S'il est un minimum expérimenté, il sera à même d'installer et configurer des pare-feux personnels sur chacun de ses équipements connectés, pour autant que cela soit possible (par exemple, les contraintes liées aux équipements de l'IoT décrites dans la sous-section 1.2.3 sont un frein à cela).

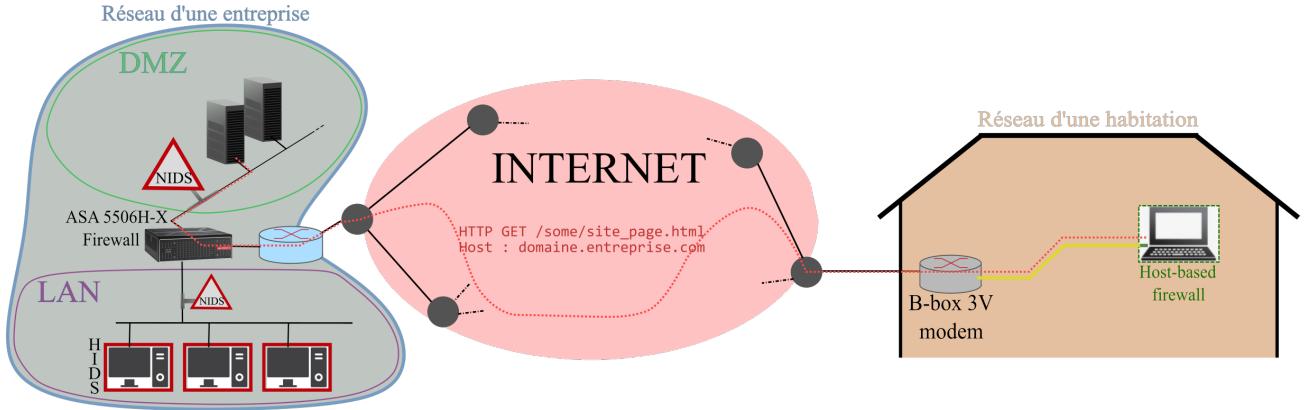


FIGURE 6 – Schéma récapitulatif des places des différents dispositifs de protection dans un réseau

2.3 Les types de pare-feu

2.3.1 Pare-feu de filtrage (sans état)

Aussi désignés dans la littérature par le terme *stateless*, ces pare-feux sont les plus rudimentaires, le premier prototype étant celui élaboré par Jeffery Mogul en 1989 [25]. Généralement, ces pare-feux sont rapides mais peu efficaces car facilement dupés. Un pare-feu de ce type inspecte chaque paquet individuellement et détermine s'il peut passer sur base d'un ensemble de règles écrites que ses en-têtes vérifient ou non. Ces règles portent sur les en-têtes TCP/IP (également UDP), les options y attenant et l'interface d'entrée. Plus en détail : les adresses IP source et destination, les ports source et destination au minimum seront soumis aux filtres du pare-feu. Il y a plusieurs problèmes avec cette approche :

- Il n'y a aucune vérification au dessus de la couche transport : la partie applicative peut contenir n'importe quoi
- Le pare-feu n'est pas dynamique : il n'apprend rien du trafic qu'il laisse passer. Aucun état n'est retenu, alors que par exemple TCP est un protocole lié à une machine à état (orienté connexion) dont il pourrait être possible de garder une trace
- Les règles sont très redondantes à écrire pour être efficaces, très peu modulables et facilement sujettes aux erreurs et oubli

Les paquets ICMP peuvent également faire l'objet d'un filtrage du même acabit [15]. Ce type de pare-feu tombe en désuétude car il est trop simple à duper par des attaquants.

2.3.2 Pare-feu à état

Le but de ce type de pare-feu dit *stateful* est d'améliorer les techniques de filtrage sans état en maintenant de l'information sur les flux passant au travers, principalement TCP. Effectivement, au contraire d'UDP, TCP est un protocole qui garde un état qui se traduit dans certains champs de ses headers. Dès lors, le pare-feu peut accéder à cette information, l'analyser et l'utiliser pour en déduire dans quel état est actuellement la connexion entre les deux processus communiquant sur les deux hôtes impliqués. Cela peut se traduire par l'inspection des flags (SYN, FIN, ...), des numéros de séquences et acquittements qui permettent au pare-feu de contrôler en conséquence le contenu d'une table interne des connexions.

L'établissement d'une connexion TCP se fait comme suit : lors de la réception du premier paquet TCP par le pare-feu, celui-ci va le soumettre à ses règles de filtrage définies comme dans un pare-feu sans état. Si le paquet est valide, il est retransmis pour poursuivre librement sa route et une entrée est ajoutée dans la table interne du pare-feu. Celle-ci va changer d'état à chaque paquet de la communication reçu, jusqu'à arriver dans un état dit *established* correspondant à la fin du 3-way handshake TCP. Tous les paquets suivants pourront alors être traités rapidement en établissant la correspondance entre ses en-têtes et celles de l'entrée en état established dans la table (qui sont donc théoriquement également légitimes). Cette façon de procéder est plus efficace que de soumettre chaque paquet individuellement aux règles de filtrage qui peuvent être complexes et demander beaucoup de travail et donc gourmandes en temps.

Dans le cas d'une communication utilisant UDP, le contrôle possible est moins fin car il ne s'agit pas d'un protocole à état et les échanges ne sont pas bidirectionnels. Dans ce cas, lorsque le premier paquet d'une transmission arrive en entrée du pare-feu, il est soumis aux règles de filtrage. S'il est considéré comme correct, le pare-feu ajoute une entrée dans sa table pour les couples d'adresses et ports et considérera les suivants comme légitimes. Cette entrée expirera après un temps donné sans autre nouveau paquet la matchant.

2.3.3 Pare-feu applicatif

Ce type de pare-feu peut être considéré comme une extension complète aux simples pare-feux à états dans le but d'en améliorer la fiabilité. Là où ces derniers sont capables de déterminer quels protocoles sont utilisés et autorisés sur chaque port, les filtres ajoutés au niveau applicatif peuvent en plus déduire à quelles fins sont utilisés ces protocoles (l'inspection va jusqu'à la couche 7 du modèle OSI). Il s'agit cependant de plus que de simples filtres, c'est une technologie à part entière qui est utilisée : l'inspection profonde des paquets ou *deep packet inspection* (DPI). Les pare-feux utilisant la DPI mêlent les fonctionnalités d'un pare-feu à état avec les systèmes de détection et prévention d'intrusions (sous-sections 2.4 et 2.5). La décision de bloquer ou de laisser passer le paquet tient compte de ce qu'il contient réellement au niveau données et de l'interprétation qui en est faite pour déterminer s'il représente un danger.

Par exemple, les *Web Application Firewalls* (WAF) sont une sous-catégorie des pare-feux applicatifs qui opèrent un filtrage sur le trafic HTTP. Ils peuvent donc filtrer le trafic sur base de certaines règles d'accessibilités fixées en bloquant des requêtes aux URL douteuses ou considérées comme non éthiques vis-à-vis de l'organisme maintenant le réseau par lequel elles transitent. Outre cela, les WAFs permettent également de se protéger contre les attaques dites de type *parameter tampering* (injection SQL, cross-site scripting, etc.) [24], ainsi que d'inclure des filtres anti-spam pour les courriels [8].

Plus généralement, les pare-feux applicatifs permettent de se protéger des virus, vers et tentatives d'*exploits* de faiblesses connues des systèmes à protéger [2] [24]. Un autre avantage d'analyser jusqu'à la couche 7 grâce à la DPI est que certains protocoles (FTP par exemple) utilisent la couche applicative pour transmettre des informations relatives aux couches plus basses (adresses IPs, ports, etc.) et qui seront impliquées dans des communications passant par le dispositif pare-feu [8]. Comme illustré par l'exemple des WAFs, le principal défaut inhérent à leur nature réside dans le fait que pour chaque type de trafic applicatif que l'on souhaite contrôler, un travail spécifique doit être opéré (établir des règles dédiées par exemple) [5].

2.3.4 Pare-feu applicatif proxy

Ces pare-feux mêlent l'inspection au niveau applicatif et le rôle de proxy pour les flux qui sont destinés à le traverser originellement. Ainsi, il devient réellement un intermédiaire entre les deux parties communiquant, interceptant les paquets pour les analyser puis les retransmettant vers le destinataire comme s'il en était la source [15]. Avant cette retransmission, le paquet peut être inspecté du point de vue de toutes ses couches et déterminé comme étant dangereux en fonction des opérations de filtrages effectuées dans le proxy. Un contrôle très fin peut donc être opéré. Cela garantit un niveau de sécurité plus haut que les pare-feux de filtrage pur qui traitent les paquets à la volée [7].

Cette approche possède cependant trois gros désavantages :

- Clairement, l'impact sur les performances du réseau n'est pas insignifiant (surtout si le pare-feu n'est basé que sur du software pur et pas sur une machine adaptée)
- À chaque nouveau protocole applicatif que l'on souhaite pouvoir filtrer de la sorte, un nouveau dispositif proxy correspondant doit être pensé et développé pour en inspecter le trafic
- Si un proxy présente une faille de sécurité, un assaillant peut l'utiliser comme un vecteur d'attaque et prendre le contrôle du système sur lequel il tourne. S'il parvient alors à désactiver les services relatifs au pare-feu, la voie vers ce qu'on souhaitait protéger se retrouve toute ouverte

2.3.5 Pare-feu identifiant

Pour ce type de dispositif, on veut pouvoir définir des règles de filtrage en fonction de l'utilisateur ou le groupe qui se cache derrière un paquet ou un flux. Plusieurs schémas d'association existent, définissant ce à quoi la notion d'un utilisateur correspond. Par exemple certains de ces pare-feux comme `authpf`² sous OpenBSD, qui étant placés sur un gateway, exigent qu'avant d'autoriser un flux l'utilisateur d'un hôte communiquant doit établir une connexion SSH avec ce dernier. Un autre type d'association possible peut se faire par adresse MAC.

Les pare-feux destinés aux entreprises développés par Cisco (ASA) [13] et Palo Alto Networks [16] utilisent des gestionnaires d'annuaires d'utilisateurs basés sur les services `Windows Active Directory`. Dans ces bases de données utilisateurs, on retrouve les différentes adresses IP associées à un utilisateur particulier ainsi que les groupes qu'ils composent. Les interactions avec le pare-feu sont transparentes à ce niveau, en effet ce dernier travaille avec les objets « utilisateurs » et non plus directement les IPs (bien que cela reste possible). Procéder de cette façon apporte plusieurs avantages [13] :

- Cela simplifie la création et la gestion des politiques de sécurité du pare-feu
- Offre la possibilité d'identifier facilement les utilisateurs utilisant le réseau
- Simplifie la surveillance des activités des utilisateurs du réseau et permet d'identifier la source de menaces

2. <http://www.openbsd.org/faq/pf/authpf.html>

2.4 Systèmes de détection d'intrusion (*IDS*)

Ce type de systèmes, abrégé par *IDS* pour *intrusion detection system*, est analogue à ce qu'est un système d'alarme intérieur dans l'environnement d'une habitation : quand les dispositifs de protection mis en place à l'entrée sont contournés (une présence non permise est détectée à l'intérieur), l'alarme est lancée et des actions sont éventuellement prises en conséquence. Les *IDSs* sont donc des sentinelles qui surveillent le réseau interne, logiquement placées après le pare-feu dans le sens entrant. Ces systèmes se présentent sous la forme d'outils spécialisés dans l'interprétation des logs des routeurs, pare-feux, serveurs et autres agents du réseau interne. Les *IDSs* sont épaulés par une base de données des signatures d'attaques déjà connues et y comparent le contenu des logs afin de trouver des *patterns* qui correspondent [15]. Dans une telle situation, plusieurs actions allant de simplement alerter l'administrateur réseau à couper les accès réseaux des machines peuvent être déclenchées, en fonction du degré de certitude (correspondance forte entre les signatures) et de la menace.

Tout comme les pare-feux, les *IDSs* peuvent être des ressources software ou reposent sur du hardware spécifique. Dans le cas où il s'agit de software, ils sont établis sur la même machine que le pare-feu, les proxy ou autres dispositifs de bordure de réseau. S'ils se présentent sous forme d'équipements hardware spécifiques, ils sont installés de sorte à contrôler et surveiller de près un de ces dispositifs sensibles. Le trafic tant entrant que sortant peut être analysé par les *IDSs*, car les attaques peuvent autant venir de l'extérieur que se déployer depuis l'intérieur (chevaux de Troie, spywares, etc.).

L'analyse des événements repose communément sur deux techniques. La première, la détection par signature (*signature detection*), utilise une base de données de signatures d'attaques déjà connues et se base sur le trafic et les patterns observés pour établir un match avec ces signatures. La seconde approche est celle de la détection d'anomalies (*anomaly detection*). Il est question d'utiliser des heuristiques afin de distinguer les situations et comportements anormaux, sur base de profils types construits par analyses statistiques, agencements de règles ou réseaux neuronaux. Outre la technique utilisée, trois catégories d'*IDSs* existent en fonction de leur place dans l'architecture réseau (leur mode de fonctionnement diffère en conséquence) :

- Surveillance du réseau (*NIDS* : *Network-based IDS*)
- Surveillance des systèmes du réseau (*HIDS* : *Host-based IDS*)
- Surveillance distribuée du réseau (*NIDS* : *Distributed IDS*)

2.4.1 Les NIDSs

Ces dispositifs surveillent le réseau ou un segment de ce réseau, sous forme d'un équipement intermédiaire de capture (ses interfaces sont en mode écoute pour l'entièreté du trafic transitant par le segment). Il est important d'avoir plusieurs unités de surveillance distinctes dans le cas où le réseau est scindé en plusieurs modules/zones. Par exemple, un serveur web d'une entreprise pourrait être infecté et servir de plateforme de lancement d'une attaque depuis le réseau interne, vers un autre module contenant des serveurs internes.

Deux manières de procéder à l'analyse sont possibles : soit en mode *in-line* soit en mode *off-line*. Le cas de la capture du trafic par interface décrit ci-dessus correspond au mode *in-line*. Le trafic est analysé en temps réel, ce qui permet une plus grande réactivité mais peut constituer un goulet d'étranglement si le processus de décision est coûteux. Le mode *off-line* est donc plus avantageux à ce niveau, puisque le processus d'analyse et décision est opéré sur des données stockées et non pas à traiter à la volée. Une inspection plus fine est alors envisageable, cependant le principal défaut de ce mode de fonctionnement reste le manque de réactivité à une attaque.

2.4.2 Les HIDSs

Les IDSs orientés systèmes, *Host-based IDS*, sont élaborés dans le but de protéger uniquement l'hôte sur lequel ils sont déployés et non plus un segment du réseau. Ils ciblent également de façon précise le trafic qu'ils doivent surveiller en fonction de la nature du système. Par exemple, si l'hôte ne maintient aucun service DNS, il est inutile d'analyser des requêtes DNS qui lui parviendraient pour y déceler une menace exploitant une faille connue dans le protocole DNS. Puisque les HIDSs vont s'exécuter sous forme d'un processus (*daemon*) sur la machine hôte qui elle-même fonctionne sous un OS classique, un lien doit être établi entre les deux pour que l'HIDS puisse surveiller le système et les interfaces réseaux. Certains vont même jusqu'à rechercher les intrusions dans le noyau de l'OS.

D'une part, un HIDS peut surveiller le comportement du système dynamiquement en récupérant des informations que ce dernier met à sa disposition à la manière d'un antivirus. Différents indicateurs sont à interpréter :

- Activité de la machine même : processus qui y vivent, ressources qu'ils consomment (CPU, RAM, réseau etc.), modification dans les comptes utilisateurs
- Activité des utilisateurs : commandes entrées, programmes lancés, tentatives d'accès à des ressources non autorisées, passage au compte administrateur
- Patterns d'exécution ou de procédure de déploiement des vers, virus, chevaux de Troie (shell ouvert simultanément à l'ouverture d'un fichier, accès anormaux aux interfaces réseau, ...)

D'autre part, certaines sections plus critiques du système peuvent être intéressantes à analyser finement, car c'est généralement là qu'un attaquant voulant prendre le contrôle de la machine va laisser des traces en y installant ses outils softwares. Un HIDS voulant s'assurer que des sections critiques (système de fichiers, registres, ...) ne sont pas infectées va maintenir une base de données de leur évolution dans le temps, sous forme d'attributs et sommes de contrôles (*checksums*). Ces dernières permettent d'assurer l'intégrité des ressources concernées en en comparant les valeurs régulièrement dans le temps, et notifiant les différences observées.

2.4.3 Les DIDSs

Les DIDSs se présentent sous la forme d'une architecture distribuée d'IDSs (surveillant des segments de réseaux et/ou hôtes), et d'une unité centrale de management qui récolte toutes les informations de ces dispositifs déployés. Ainsi, la machine centrale peut maintenir une vaste base de données centralisée, représentant l'état global du réseau à protéger. Cette approche présente plusieurs avantages :

- moins de faux positifs car plus de données pour justifier une prise de décision (agrégation des événements et attaques)
- mises à jour et distribution de la base de données des signatures aisées
- centralisation des alertes/logs, contrôle global avec une vue d'ensemble
- réponses aux événements plus efficaces, IDSs plus simples à administrer en conséquence (ajout de nouvelles règles suite à des brèches découvertes, blacklisting d'IPs, etc.)

Afin de transmettre les alertes générées jusqu'à l'unité centrale efficacement, les différents nœuds IDSs du réseau devraient communiquer de façon homogène. À cette fin, un standard a été écrit sous le nom de format IDMEF (*Intrusion Detection Message Exchange Format*, RFC 4765). Ce langage d'alertes est lisible par l'humain, utilisant le format XML. Utilisé par tous les nœuds qui effectuent la tâche de détection d'intrusion (quelque soit le type et l'implémentation), l'unité centrale qui récolte toutes les alertes peut utiliser un outil de pilotage de la sécurité du réseau global comme [Prelude](#)³. Cet outil normalise, trie, agrège, corrèle et en tire les conclusions sur des décisions à prendre pour protéger le réseau, tout en fournissant une interface (et des logs) via laquelle l'administrateur peut surveiller ce qui se passe globalement.

3. [https://en.wikipedia.org/wiki/Prelude_SIEM_\(Intrusion_Detection_System\)](https://en.wikipedia.org/wiki/Prelude_SIEM_(Intrusion_Detection_System))

2.5 Systèmes de prévention d'intrusion (IPS)

Dans la section précédente, il était question de détection des attaques et pour certaines d'entre elles des actions à effectuer en contre-mesure pour bloquer la menace. Cela relève du domaine d'application des systèmes de prévention d'intrusion, aussi abrégé IPS pour *intrusion prevention systems*. Tout comme pour les IDSs, deux types d'IPS peuvent être distingués en fonction de leur place dans l'architecture réseau : les *host-based IPS* agissent au niveau local pour un hôte, tandis que les *network-based IPS* défendent un réseau ou un de ses segments internes. De fait, IDSs et IPSs vont de paire puisque les contre-mesures prises par un dispositif IPS sont basées sur ce que l'IDS détecte et lui transmet comme information.

En plus de lancer une alerte à destination d'un administrateur et écrire des logs complets, les actions qu'un IPS peut effectuer suite à la détection d'une menace sont généralement les suivantes :

- Jeter les paquets détectés comme contribuant à cette attaque
- Bloquer tout trafic issu de la même adresse IP/du même utilisateur que celui considéré comme attaquant (mettre fin à la connexion TCP impliquée)
- Reconfigurer le pare-feu associé pour qu'il puisse à l'avenir bloquer le trafic relatif à cette attaque
- Si l'IPS est lié à un dispositif faisant du proxy, il peut agir sur le contenu des paquets pour neutraliser la menace (par exemple ôter un fichier infecté joint à un mail)

2.6 Les pare-feux nouvelle génération

Abrégés par NGFW (*next-generation firewalls*), ces pare-feux font suite à ce qu'on distingue comme la troisième génération dans l'évolution des dispositifs pare-feux [5], c'est-à-dire ceux qui appliquent les concepts de filtrage au niveau applicatif décrits dans la sous-section 2.3.3. Les NGFW effectuent une analyse plus fine que ces derniers en utilisant les technologies avancées de *deep packet inspection*. Les services qu'offrent la nouvelle génération de pare-feux s'étendent encore au-delà de la simple fonctionnalité d'inspection des paquets, et sont devenus des systèmes très complexes implémentant d'autres dispositifs liés à la sécurité dont les suivants :

- Support des technologies de (dé)chiffrement utilisées classiquement dans les réseaux
- Les systèmes de management et contrôle sur base des identités des utilisateurs inspirés des pare-feux identifiants
- Les pare-feux destinés aux flux web (WAFs), le filtrage d'URLs selon des politiques de sécurité et éthique
- Des moyens de contre-mesures aux attaques empruntés aux IPSs
- Les environnements virtuels (*sandboxing*) de tests pour déceler dans un flux suspect une attaque non reconnue [16]
- De vastes bases de données sur le cloud interrogées efficacement via un service spécialisé fourni avec le pare-feu (environnement WildFire et similaires)

2.7 Les pare-feux de référence

2.7.1 Sous forme de software

Ces pare-feux sont destinés à être installés et utilisés par-dessus des systèmes d'exploitation génériques tels que Windows, distributions Linux et BSD, etc. En plus de protéger la machine sur laquelle ils opèrent des menaces extérieures, certains d'entre eux intègrent d'autres fonctionnalités telles que du *traffic shaping* ou la gestion d'un système de NAT.

Windows

Windows Firewall est intégré par défaut avec les versions de l'OS Windows postérieures à *Windows XP*. Il s'agit d'un pare-feu qui se veut « user-friendly » : par défaut il présente une liste de profils pré-configurés pour que l'utilisateur puisse sélectionner celui qui convient le plus à ses besoins (*Public*, *Private* et *Domain*). Il agit comme un pare-feu à état basique mais assez complet dans sa fonction restreinte cependant, en plus d'être simple à utiliser et de disposer d'une interface graphique. Le fait qu'il soit directement incorporé dans le système permet un contrôle fin de quelles applications sont autorisées pour quels types de trafics.

Linux

Netfilter est la solution de base de filtrage de paquets inclue dans le noyau Linux. Plusieurs modules liés au noyau sont également fournis pour servir d'intermédiaire dans la configuration de Netfilter, notamment sous la forme de tables de règles en chaînes à appliquer séquentiellement sur les entrées à filtrer. Ces modules sont accessibles depuis l'utilitaire *iptables*, sous les noms de *ip_tables*, *ip6_tables*, etc.

nftables a pour objectif de remplacer Netfilter à terme car il est sensé apporter plus de performances principalement. L'utilitaire permettant de faire le lien avec cette nouvelle solution implémentée dans le noyau est réduit à *nft* (réduction du nombre de module par rapport à Netfilter).

Shorewall repose sur Netfilter, faisant office de couche d'abstraction pour faciliter la configuration des règles qui est assez laborieuse avec les outils comme *iptables*. Il permet l'écriture des règles alimentant les tables par un mécanisme de fichiers de configuration. C'est un software complet, il propose des fonctionnalités classiques telles que du NAT, du *traffic shaping*, un support VPN et **autres**. Une autre fonctionnalité notable est la possibilité de partitionner le réseau en zones distinctes (ou pouvant se recouvrir l'une l'autre), permettant le contrôle de toutes les connexions entre chaque paire de zone à considérer. Cela permet notamment l'établissement aisément de zones démilitarisées (DMZ).

BSD

Packet Filter (pf) est comparable à Netfilter pour les noyaux Linux. C'est une solution de filtrage de paquets à état, incluant d'autres fonctionnalités classiques telles que la gestion de NAT, des mécanismes liés à la *qualité de service* (QoS), une gestion des logs avancée en offrant leur configuration règle par règle. D'autres utilitaires sont venus s'y greffer, tel que *authpf* qui permet un système de gateway authentifié (voir les pare-feux identifiants). Packet Filter a fait l'objet de port vers de nombreux autres systèmes d'exploitation, notamment Apple Mac OS X et iOS, NetBSD et Solaris.

2.7.2 Sous forme d'appliance

Le terme *appliance* désigne un appareil/système constitué du strict minimum de composants pour remplir au mieux une fonction spécifique. Du point de vue des dispositifs pare-feux, on peut distinguer trois catégories d'appliances :

- les pare-feux destinés aux systèmes embarqués, particulièrement contraints en ressources comme c'est le cas pour certains équipements de l'IoT.
- les pare-feux logiciels destinés à fonctionner au dessus d'une couche d'abstraction du hardware spécifique (ou même sur une machine virtuelle), basés sur un noyau d'OS qui est transparent à l'utilisation
- les pare-feux reposant sur du hardware (machines dédiées vendues telles quelles) et qui utilisent le matériel et les équipements spécifiques construits à cette fin apportant de meilleures performances

Généralement, il s'agit de solutions destinées aux entreprises de par leur prix et leur nature. Effectivement, comme ce sont des équipements/dispositifs dédiés (presque) uniquement à la fonction de pare-feu (et à la sécurisation du réseau en général), ils sont souvent déployés comme des pare-feux niveau réseau. À noter que beaucoup de ces pare-feux tournent sur un OS propriétaire de l'entreprise qui les fabrique (Cisco, Check Point, etc.). Cela permet aux fabricants d'optimiser leurs produits tant au niveau performance que sécurité, restreignant leur OS au strict nécessaire (certains d'entre eux sont basés sur des noyaux déjà existants).

Appliances sur systèmes embarqués

IPFire est dérivé du noyau Linux et est capable de remplir toutes les fonctions basiques d'un routeur utilisé en conjugaison avec un pare-feu, sous des contraintes de ressource de 1GHz pour le CPU, 1GB de RAM et 4GB d'espace disque. De fait, IPFire peut être utilisé sur des systèmes comme une carte Raspberry PI ou dans des environnements virtuels. IPFire est modulable avec des [add-ons](#) gérés par un utilitaire de management, *Pakfire*. Cependant, [la version minimaliste](#) supporte déjà des fonctionnalités telles que l'inspection des paquets à état, le proxying, la prévention d'intrusions, la segmentation en zones du réseau, la gestion de QoS, les services DHCP et DNS, etc.

IPCop est un pare-feu open-source assez minimalistique utilisant Netfilter en arrière-plan. Les pré-requis pour le faire tourner sont : 200MHz de fréquence CPU, 64 Mo de RAM et 800 Mo d'espace de stockage. Même avec cette configuration minimale, il assure les fonctionnalités de proxy et d'IPS en plus du filtrage à état classique. Des modules peuvent être ajoutés en fonction des services supplémentaires requis : services DHCP ou DNS, *traffic shaping*, liste d'accès utilisateurs, etc. IPCop est une branche divergente du projet [SmoothWall](#), ce dernier s'étant développé de telle sorte qu'il n'est plus associable à un pare-feu *embarqué* en plus d'être passé en partie sous licence commerciale.

Appliances orientées software

pfSense se repose sur FreeBSD (bien qu'aucune connaissance sur ce dernier ne soit requise, puisque pfSense est une appliance à utiliser tel quel). Le software peut être installé sur une machine physique dédiée ou dans un environnement virtuel. Il est complètement open-source, et [à la prétention](#) d'offrir tous les services que les grandes distributions incluent dans leur appliances physiques en un seul software, gratuit et extensible. pfSense opère comme un routeur et pare-feu sur la machine dédiée, ces services étant alors configurables et mis à jour via une interface web.

Untangle est quant à lui basé sur le noyau Linux. Il est également déployable sur une machine simple dédiée ou virtuelle. Ses [fonctionnalités](#) incluent entre autres les classiques d'un pare-feu nouvelle génération (filtrage du contenu, IPS, identification utilisateur, etc.) mais aussi la possibilité de faire du cache web, du *load balancing* et du routage. Cependant, seule une version légère du software est distribuée gratuitement.

Softwares reposant sur des appliances hardware

CheckPoint Software est une entreprise qui vend des machines dédiées à la protection des réseaux ainsi que les solutions logicielles qui les accompagnent. Le software *GAiA* fait office d'OS pour les [appliances hardwares vendues par CheckPoint](#) pour lesquels il est optimisé, mais peut également être déployé sur une appliance virtuelle compatible hébergée sur une [machine qui en est capable](#).

Cisco offre des services similaires avec ses pare-feux nouvelle génération ASA 5500-X. Le software Cisco *Adaptative Security Appliance* est l'OS dédié principalement aux machines de la famille ASA (encore une fois des solutions de virtualisation existent). La sur-couche *Cisco Firepower Threat Defence* (FTD) est un software intégrant les fonctionnalités d'ASA et des services axés IDS/IPS nouvelle génération de [FirePOWER](#).

Ces solutions sont généralement utilisées dans des réseaux de grande taille, donc en entreprise. Les appliances hardwares sont vendues telles quelles à prix élevé, et le software fourni avec y est intégré pour en garantir les performances, tant au niveau des débits que de la détection et du blocage des menaces. Cependant, il n'est pas vain de s'y intéresser car certaines technologies sur lesquelles elles sont basées sont parfois réutilisables à plus petite échelle (dans le cadre de solutions pour protéger un réseau domestique entre autres). Par exemple, FirePOWER a été développé à partir du software open-source [Snort](#) et utilise les bases de données de définitions de signatures virales issues d'un utilitaire anti-virus nommé [ClamWin](#), libre également.

3 La sécurité dans l'IoT

3.1 Motivations et exemples

À priori, dans le cadre d'une smarthome, les équipements relevant de l'IoT sont plutôt des petits accessoires, voire gadgets (montres, lampes, balances connectées par exemple). On pourrait donc se dire qu'introduire des mécanismes de sécurité avancés dans de tels petits systèmes n'est pas primordial, l'aspect fonctionnel étant plus mis à l'honneur. C'est d'ailleurs ce qu'ont fait beaucoup de constructeurs, mettant sur le marché des équipements possédant des failles de sécurité importantes ou des faiblesses au niveau du design de conception. Un exemple interpellant est le fait que pour un équipement qui possède une interface accessible par un portail de sécurité par identification (requérant une entrée de la forme *user/password*), les valeurs d'usine par défaut sont laissées telles quelles, l'utilisateur n'étant jamais invité à les modifier [20]. Ainsi, il existe des listes établies de ces identifiants par défaut en fonction du constructeur ([pour les webcams](#) par exemple) qui, combinées avec des outils puissants comme [Shodan](#) ou [Censys](#), constituent des brèches dans la protection de l'habitation pour les attaquants qui savent les exploiter. La Figure 7 illustre la déconcertante facilité avec laquelle il est possible d'attenter à la vie privée des utilisateurs inconscients (une unique requête a été nécessaire).

IP Webcam

86.205.251.170
amuseille-552-1-18-170.wf6-205.abo.wanadoo.fr
Orange
Added on 2018-12-16 20:32:02 GMT
France, Mâlemort
Technologie: swf [B](#) [F](#)
[Details](#)



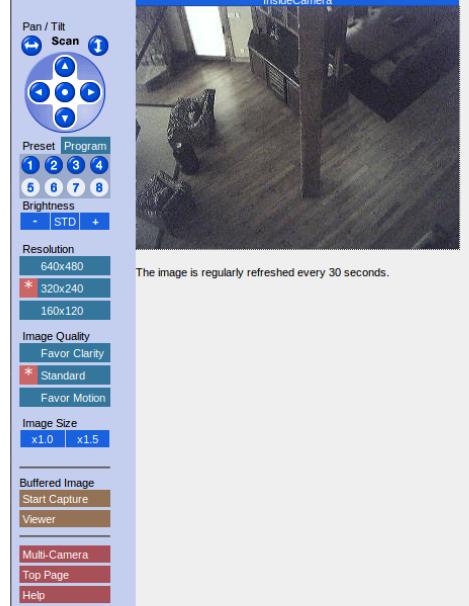
```
HTTP/1.1 200 OK
Connection: close
Server: IP Webcam Server 0.3
Cache-Control: no-store, no-cache, must-revalidate, pre-check=0, post-check=0, max-age=0
Pragma: no-cache
Expires: -1
Access-Control-Allow-Origin: *
Content-Type: text/html
```

IP Webcam

107.181.14.128
Phoenix Cable
Added on 2018-11-26 10:05:11 GMT
United States, Phoenix City
[Details](#)



```
HTTP/1.1 200 OK
Connection: close
Server: IP Webcam Server 0.4
Cache-Control: no-store, no-cache, must-revalidate, pre-check=0, post-check=0, max-age=0
Pragma: no-cache
Expires: -1
Access-Control-Allow-Origin: *
Content-Type: text/html
```



The image is regularly refreshed every 30 seconds.

Android webcam server

If no video appear below, you are using unsupported browser

Tested browsers: [Mozilla Firefox](#), [Google Chrome](#)

[Open camera controls](#)



IP Webcam Remote

[Turn on LED](#) [Turn off LED](#)
[Autofocus](#)
[Cancel autofocus](#)
[Take full-res picture](#)
[Autofocus and take full-res picture](#)

Audio, in case your external player does not support separate streams:

[Click here to play audio with browser](#)
[Click here to play audio in external media player. Right click here and select "save as" to record it.](#)
[Fullscreen view](#) (click video to change aspect behavior, F11 to remove browser borders).

FIGURE 7 – Résultats d'une brève recherche avec Shodan ciblant des caméras en réseau non ou peu protégées

Évidemment, l'utilisateur « final » n'est pas le seul qui puisse en subir les conséquences. Ses équipements peuvent entre autres rejoindre un ensemble d'autres objets connectés infectés, reconfigurés malicieusement en vue d'effectuer une attaque de masse de type *DDOS* sur une cible étrangère au réseau domestique. C'est ce qui s'est passé avec *Mirai*, un malware catégorisé comme *botnet* décelé bien après son expansion et qui a été utilisé dans plusieurs attaques menées sur le fournisseur de services DNS *Dyn*, la société d'hébergement *OVH* et bien d'autres. Mirai n'est pas un cas isolé, et les hackers ont pleinement saisi l'opportunité qu'incarne l'IoT : il existe beaucoup de failles connues et exploitables sur des équipements hétérogènes, et celles-ci sont en partie rendues publiques sur par exemple des dépôts *githubs* et des forums dédiés.

Les équipements qu'on retrouve dans une smarthome ne sont pas les seuls à présenter des lacunes en terme de sécurité. Encore plus grave, les systèmes connectés utilisés dans l'industrie de nos jours peuvent être attaqués, et cela peut porter lourdement à conséquence (jusqu'à l'échelle internationale). Le scénario *Horus* le démontre particulièrement bien. Imaginée par un chercheur néerlandais, cette attaque cible les installations photo-voltaïques qui fournissent en énergie l'Europe, plus spécifiquement les onduleurs connectés y opérant (qui, brièvement, sont les dispositifs transformant le courant alternatif en courant continu). Pas moins de 17 failles ont été découvertes, dont certaines permettant de prendre le contrôle total des onduleurs à distance. Il serait alors possible pour un attaquant de simuler artificiellement l'impact qu'une éclipse solaire aurait sur le réseau, qui est en temps normal compensé artificiellement (celle-ci pouvant être anticipée). En appliquant une telle variation de puissance soudaine, le réseau entrerait dans un état de fluctuation tel qu'il ne serait plus contrôlable et une grande partie du réseau finirait par sauter. La perte suite à une telle panne, considérant qu'elle dure trois heures, a été estimée à 4,5 milliards d'euros.

3.2 Description des besoins en terme de sécurité

Ces notions sont similaires à celles généralement mises en avant dans le domaine de la sécurité informatique en général. On les agrémentera en faisant le lien avec le cas un peu plus restreint des équipements relatifs à l'IoT, si possible. Ces grands axes sont les suivants [19] [21] :

Confidentialité

L'accès aux informations doit être limité aux personnes et systèmes qui y sont autorisés, et uniquement eux : aucun accès indésirable ne doit être possible. Dans un réseau constitué de plusieurs équipements, les communications entre chacun des nœuds doivent être protégées des intrus pouvant potentiellement les intercepter.

Intégrité

Les informations stockées et échangées ne doivent pas être altérées, quelle qu'en soit la raison (de source malveillante ou non). En plus des informations, le firmware exploitant les équipements doit également respecter cette propriété (le système ne doit pas être compromis).

Disponibilité et continuité

Les systèmes doivent être fonctionnels quand cela est attendu et garantir l'accès et la bonne exécution des services qu'ils sont destinés à offrir, dans le délai attendu. La sécurité doit pouvoir être assurée en cas de panne ou dysfonctionnement de certaines parties du système, et ce même dans le cas de systèmes dont les ressources sont limitées (équipements sur batterie en fin de vie par exemple).

Authentification

Seuls les utilisateurs légitimes (acquéreurs du produit entre autres) devraient être en mesure d'accéder aux systèmes et de les configurer, ainsi que d'accéder aux informations sensibles qu'ils possèdent. L'authentification peut aussi être relative à une entité, comme le fabricant certifié du produit qui tente d'atteindre le système pour effectuer sa mise-à-jour ou lui appliquer un patch afin de corriger une faille dans sa sécurité.

3.3 Vulnérabilités liées aux équipements IoT

Les caractéristiques généralement présentées par les équipements décrites comme des restrictions dans la sous-section 1.2.3 portent à conséquence en terme de sécurité. Il en résulte de potentielles vulnérabilités [21], couvertes ou non par ces équipements en fonction de leur conception et implémentation spécifique. Celles-ci sont brièvement décrites ci-dessous.

L'exemple des caméras IP présenté en introduction l'illustre bien : l'**accessibilité aux équipements depuis le réseau** (surtout Internet) est un problème conséquent, dans la mesure où il n'est pas pris en considération à la conception du système. Des mesures découlant du principe d'authentification doivent être prises afin de limiter les accès aux entités légitimes uniquement.

Une autre vulnérabilité découle directement du fait que les équipements sont **restreints en ressources CPU et mémoire**. On y retrouve des petits microcontrôleurs travaillant sur 8 bits [21], ce qui contraint beaucoup l'utilisation des algorithmes de chiffrement classiques (AES, 3DES, TEA, etc.). C'est donc une vulnérabilité importante qu'il est capital de contrer car transmettre des données en clair va à l'encontre des principes de confidentialité énoncés plus tôt. Des solutions sous forme d'algorithmes adaptés [11] ou de matériel spécialement dédié au chiffrement⁴ existent.

Outre les restrictions en terme de capacité, une autre composante est à prendre en compte : **la source d'énergie limitée** (pour les équipements utilisant une batterie par exemple). Cette vulnérabilité ouvre la porte aux attaques d'exhaustion, consistant à désactiver un équipement par le biais de l'épuisement de sa source d'alimentation [22] (attaque dite de *Depletion-of-Battery*). C'est une atteinte au principe de disponibilité.

Les **firmwares non mis à jour** et globalement les **constructeurs peu impliqués dans la sécurité de leurs produits** constituent des vulnérabilités qui peuvent être exploitées par des attaquants. Par exemple, les identifiants par défaut ou stockés en clair. Peu d'équipements installés dans les smarthomes bénéficient d'un service de mise à jour permettant des patchs de sécurité [21] suite à des failles décelées et rendues accessibles publiquement.

L'**accès physique au(x) média(s) de transmission** représente une autre vulnérabilité, surtout dans le cadre des communications sans-fil. Dans l'environnement d'une smarthome par exemple, un dispositif de sécurité installé pour protéger l'habitation qui utilise du sans-fil devrait tenir compte du fait que toutes les communications peuvent être entendues en dehors de l'enceinte des murs de la maison (par des *outils de sniffing*). En plus de l'écoute passive, le média peut aussi être pris d'assaut activement, dans l'exemple précédent une attaque de type *Denial Of Service* pourrait impacter l'efficacité du système de sécurité.

L'**hétérogénéité des équipements** peut être problématique pour d'une part définir des standards et d'autre part avoir une bonne connaissance technique de chacun des équipements évoluant par exemple dans une smarthome. La documentation pour chaque équipement devrait alors lui être spécifique, et pour peu qu'elle existe il faudrait qu'elle développe les mécanismes de sécurité fournis avec l'équipement. Allant de paire avec cela, les systèmes de mises-à-jours et patchs de sécurité divergent de l'un à l'autre.

Dans le cas particulier des smarthomes, le **manque de connaissances des utilisateurs** en terme de sécurité informatique [23] peut être considéré comme une vulnérabilité car il peut en résulter une mauvaise installation ou configuration des systèmes menant à des failles exploitables. Conjugué avec le fait que les constructeurs des équipements que l'habitant achète ne mettent pas la priorité sur leur sécurisation, cela ouvre de multiples portes que les attaquants n'ont qu'à ouvrir pour attenter à la vie privée des utilisateurs.

4. <https://www.zyngbit.com/make-things-secure-1/>

3.4 Architecture globale des systèmes impliquant des équipements IoT

Leloglu propose dans [19] un travail de synthèse de plusieurs sources dont il dérive entre autres un modèle structurel de l'IoT. Il n'est pas à considérer du point de vue d'un seul équipement, mais plutôt au niveau global, c'est-à-dire tout ce qui est impliqué dans le fonctionnement d'un système exploitant les services d'équipements de l'IoT. C'est une représentation en couches, une couche supérieure s'appuyant sur les services fournis par les couches plus basses afin de remplir la fonction qui lui est dédiée. On distingue les couches suivantes (reprises par des exemples dans le tableau 1 et schématisées par la figure 8) :

La couche perception

Il s'agit de la couche la plus basse, où on considère le rôle des équipements IoT par rapport à l'environnement dans lequel ils évoluent. Comme expliqué dans la sous-section 1.2.3, ces derniers sont généralement conçus pour remplir une fonction spécifique. C'est dans cette couche qu'elle prend tout son sens, car elle dépend souvent de l'environnement. L'exemple le plus évident est les capteurs, possédant ce qui est nécessaire en terme de matériel pour effectuer une mesure d'une grandeur physique. Les divers systèmes embarqués peuvent également s'intéresser à leur environnement extérieur ou au(x) dispositif(s) qu'ils supervisent (lave-vaisselle connecté, etc.). Une autre technologie que l'on situe à cette couche est celle de la radio-identification (RFID - *radio frequency identification*), considérée par certains comme les prémisses à l'IoT [12]. La perception se fait ici au niveau du lecteur RFID, qui capte et interprète les tags RFID des objets qui évoluent dans son environnement. En résumé, l'objectif de cette couche est l'identification des objets uniques et la collecte d'informations de l'environnement physique.

La couche réseau

Cette couche a pour objectif le transfert des informations récoltées par la couche perception depuis les équipements qui les ont réalisées et jusqu'au système qui en effectuera le traitement. On distingue généralement deux parties dans ce transfert, la première correspondant à l'acheminement des données à travers un réseau local jusqu'à un portail (le gateway). Celui-ci qui constituera le point de départ de la deuxième partie du transfert, utilisant un réseau non spécifique tel que Internet pour atteindre le système de traitement terminal. Afin de rendre possible les communications entre ces deux réseaux de natures différentes, il est parfois nécessaire d'introduire une couche d'adaptation dans la pile des protocoles utilisés, e.g. *6LoWPAN* qui permet des échanges basés sur IPv6 dans les réseaux contraints en ressources.

La couche support

La couche support englobe tous les systèmes de traitement des données, celles-ci ayant été acheminées jusqu'au centre de traitement par le réseau. Il s'agit donc de partir des données brutes transmises et de leur appliquer un processus de transformation. Ce dernier peut comprendre de l'agrégation de données, de l'analyse, du tri, de la classification, etc., l'objectif étant d'en tirer des données utiles et pertinentes. Les données finales peuvent ainsi être stockées dans une base de données sur laquelle une ou des applications peuvent travailler.

La couche application

Les applications sont diverses et variées, développées dans des domaines tels que repris dans la figure 1. Pour remplir leurs objectifs, elles utilisent les données obtenues en sortie du processus de traitement opéré par la couche inférieure (les couches support et application étant très proches, on les considère parfois comme une seule).

Niveau dans le système	Exemples de technologies, systèmes et équipements y opérant
Couche perception	Caméras, microphones, capteurs, GPS et divers systèmes embarqués mesurant et interagissant avec l'environnement qui leur est extérieur (montres intelligentes, détecteurs, etc.), ainsi que la technologie RFID. Côté software : OS adapté aux systèmes embarqués tels que Contiki, tinyOS, LiteOS, etc.
Couche réseau	Réseaux de capteurs, réseau mobile 2G/3G, réseaux classiques tel que Internet (IP).
Couche support	Cloud, centre de traitement des données, base de données et autres technologies et supports permettant d'agrégger, stocker et analyser les données (<i>Big Data</i>)
Couche application	Applications relatives aux smarthomes (interfaçage avec l'utilisateur, optimisation de son quotidien, etc.), applications médicales (surveillance de patients, aides aux invalides, etc.), sociales et structurelles (smartcity : parkings intelligents, etc.), scientifiques (études des mesures, détection de catastrophes, etc.) et industrielles (arrosage de plants, transport, etc.)

TABLE 1 – Architecture d'un système IoT proposée par Leloglu [19]

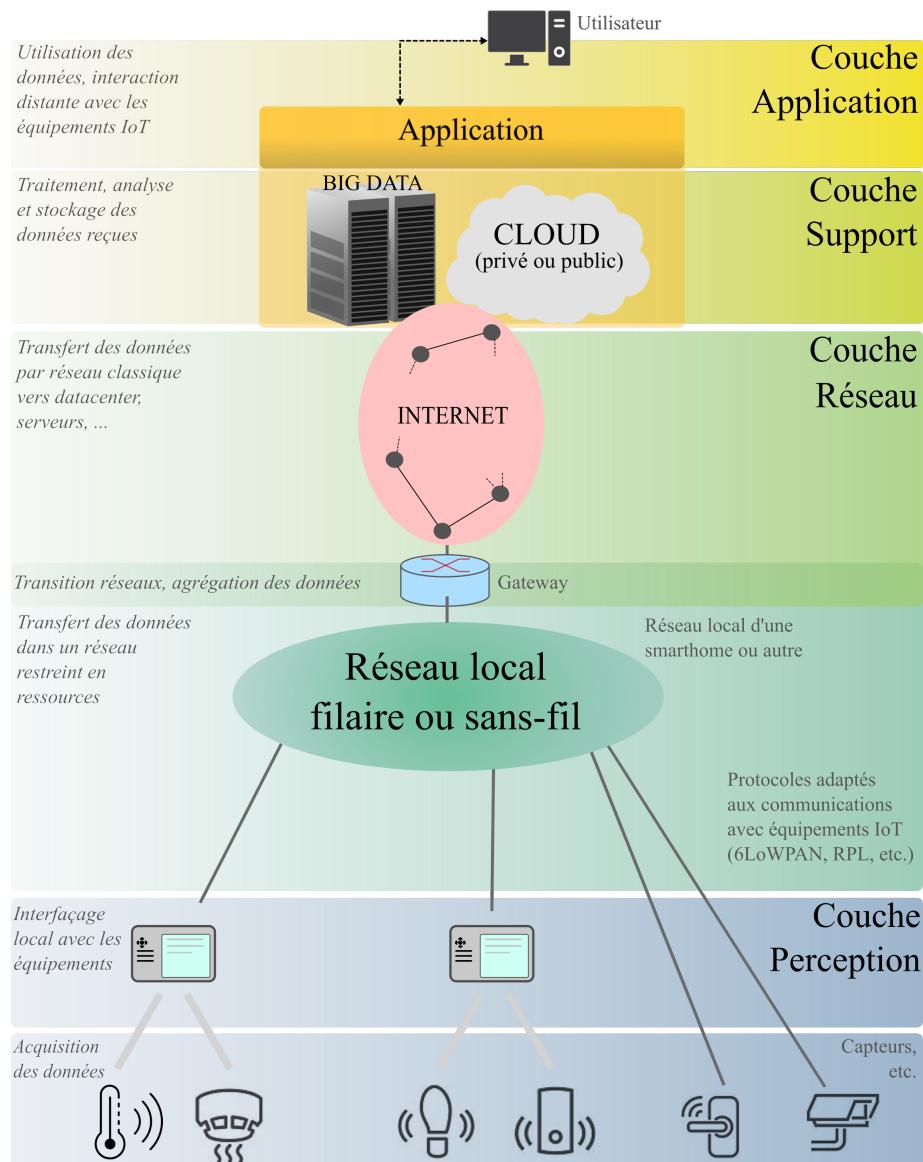


FIGURE 8 – Schématisation des couches considérées pour un système utilisant des équipements IoT

3.5 Les archétypes d'attaques courantes

Plusieurs schémas d'attaques peuvent être distingués, exploitant ou non des vulnérabilités recensées dans la section 3.3. Ces scénarios d'attaques typiques sont classés en fonction de la couche dans laquelle ils opèrent, selon Makhdoom et al. [22] et Leloglu [19].

3.5.1 Sur la couche perception

Drainage de batterie (*Depletion-of-Battery attack*) : en apparence justifiée car il s'agit d'un scénario légitime du point de vue de l'utilisation, ce type d'attaque porte préjudice à un nœud en le bombardant de requêtes qu'il doit traiter car considérées comme légitimes. Cela empêche le nœud de passer en mode veille pour préserver de l'énergie, et peut mener à l'extinction de celui-ci une fois sa source d'énergie épuisée.

Sabotage du nœud (*Node capturing/cloning*) : un attaquant peut attenter physiquement au système : il est possible qu'il accède à un nœud parmi ceux dispersés dans un réseau de capteur par exemple. Il lui est donc potentiellement possible de le corrompre ou le remplacer physiquement par un nœud intrus qui effectuera un travail dissimulé.

Brouillage radio (*Signal/Radio jamming*) : un type d'attaque DoS qui vise à occuper le canal (sans-fil) au moment où des nœuds souhaitent également émettre, brouillant les communications.

Écoute clandestine (*Eavesdropping*) : dans le cadre des communications sans-fil, un attaquant peut placer physiquement ou corrompre un nœud alors clandestin qui lui permet d'écouter les informations passant par ce canal. Ces informations peuvent être sensibles, telles que des mots de passe transmis en clair par d'autres équipements. Outre cela, le simple fait de savoir quand des informations sont émises sur le canal peut profiter à un attaquant, par exemple en l'a aidant à évaluer quand une maison est occupée ou non.

Attaque par intrusion ((*Semi-)**Invasise intrusions*)⁵ : par l'étude des variations de certains indicateurs tels que le voltage, le signal de l'horloge, etc., un attaquant peut inférer les opérations qu'un microcontrôleur effectue. Il est aussi possible de déduire certaines valeurs stockées dans la mémoire sensée rester privée telles que les clés maîtres de chiffrement (AES ou autre). En appliquant des fluctuations physiques, il est aussi possible de faire en sorte que le CPU exécute au final une autre opération que celle attendue.

Mises à jour faussées : des implémentations mal faites des mécanismes d'authentification peuvent aider un attaquant à soit maintenir une version du firmware obsolète, soit faire passer une mise à jour falsifiée comme provenant du constructeur officiel de l'équipement. Les patchs relatifs à la sécurité peuvent ainsi être bloqués, et les failles qu'ils corrigeaient laissées exploitable.

3.5.2 Sur la couche réseau

Spoofing : effectuée par le biais de paquets forgés, ce type d'attaque vise à usurper une identité renseignée comme source du paquet falsifié. Une attaque de ce type est rendue possible à plusieurs niveaux⁶ des communications, dès qu'aucun mécanisme n'est mis en place pour authentifier les entités communiquant (on parle d'*ARP*, *IP*, *MAC*, ... *spoofing*). L'usurpation d'identité permet des attaques plus concrètes telles que les *MITM attacks* décrites ci-dessous.

Selective Forwarding : un nœud infecté ou malicieux introduit dans le réseau peut influencer celui-ci quand il est question de retransmettre des paquets (il effectue un travail de routage). Selon une certaine procédure créée pour manipuler artificiellement le trafic dans un sens profitant à l'attaquant, le nœud ne retransmet que les paquets qu'il sélectionne. Afin que ce ne soit pas trop évident de déceler le subterfuge, il n'arrête pas toutes les retransmissions.

Attaque de Sybil : un nœud malicieux gère plusieurs identités simultanément, propres ou en se faisant passer pour un autre nœud. Du point de vue des nœuds légitimes, ce sont donc plusieurs équipements distincts (simultanément ou représenté tour par tour). Une panne ou une attaque impactant la disponibilité d'autres nœuds peut ainsi être masquée par ce procédé.

5. https://www.cl.cam.ac.uk/~sps32/semi-inv_def.html

6. https://en.wikipedia.org/wiki/Spoofing_attack

Sinkhole/Blackhole attack : l'attaque consiste à, à partir d'un nœud infecté, consommer les ressources des nœuds voisins dans le rayon d'émission. Cela se fait essentiellement en occupant le canal radio, brouillant les transmissions légitimes. Il en résulte une consommation d'énergie anormalement élevée pour les nœuds dans cette zone, attribuée aux retransmissions alors nécessaires.

Wormhole : cette attaque cible les protocoles de routages, introduisant une confusion dans le calcul des routes. Un nœud infecté va transmettre un paquet qu'il est sensé router, directement vers un autre nœud éloigné en établissant un tunnel invisible pour le protocole de routage. Ainsi, des nouvelles routes vont apparaître, donnant l'illusion que les deux nœuds sont proches, ce qui pour un protocole de routage se basant sur la distance entre deux nœuds constituera une nouvelle meilleure route (qui ne l'est en réalité pas du tout).

Man-in-the-Middle attack (MITM) : forme d'eavesdropping où l'attaquant s'interpose entre deux nœuds voulant communiquer légitimement. Il opère de façon à ce que la communication ait lieu normalement du point de vue des deux autres intervenants, mais de sorte que chacun communique avec lui qui joue le rôle d'intermédiaire. Il peut alors accéder au contenu de tous les échanges et le manipuler pour en tirer des informations confidentielles et influencer le déroulement de la communication.

Hello-flood attack : inondation du canal par des messages inutiles d'annonces de voisin qui demandent une réponse de la part des nœuds qui le reçoivent (générant ainsi davantage de trafic).

3.5.3 Sur la couche support et application

Altération/exploitation des données : comme ces couches manipulent des données issues de ce que les équipements ont mesuré et capté, elles peuvent être sensibles et doivent être protégées. Un attaquant qui met la main sur ces données pourrait d'une part les corrompre et d'autre part les analyser et les exploiter afin de mener une autre attaque contre les utilisateurs. Par exemple, il pourrait dégager des données enregistrées quelles sont les plages horaires pendant lesquelles un domicile n'est pas occupé et donc trouver une fenêtre de liberté pour une approche physique. De plus, ces données ne sont plus directement sous le contrôle de l'utilisateur (stockées dans les machines d'une entreprise), cela peut donc être problématique et mener à une atteinte à sa vie privée (revente illégale d'informations à une compagnie tierce, etc.).

Attaques de déni de service (*DoS, DDoS*) : les compagnies qui hébergent les services repris dans les couches support et application (traitement des données, accès aux serveurs, etc.) sont elles aussi susceptibles de se faire attaquer directement. Les attaques DoS (ou DDoS comme illustré avec Mirai en introduction de cette section) sont courantes et peuvent complètement interrompre la disponibilité de ces services par différent biais (congestion, mise à l'arrêt des machines).

Injection de code : l'attaquant utilise des failles dans les applications pour injecter et faire exécuter son propre code écrit à des fins malicieuses. Cela se fait notamment sur les bases de données (injection SQL). Une injection peut résulter en l'installation d'un malware, des requêtes illégitimes effectuées sur les bases de données, un gain de priviléges dans le système pour l'attaquant ou encore rendre possible du *cross-site scripting*⁷.

Sniffers/loggers : les attaquants récoltent un maximum des données utilisées et transitant (par exemple lors des échanges avec l'utilisateur par le biais de l'application) afin d'en tirer des informations de valeur (bien que cela soit rendu plus difficile par le chiffrement). Il peut s'agir d'identifiants, de mots de passe, d'emails, etc.

Session Hijacking : les attaquants exploitent des failles dans les processus d'authentification et gestion de session afin d'interférer dans les échanges et obtenir des accès privilégiés à des ressources. Cela se fait principalement via les cookies.

Ingénierie sociale : l'attaquant extrait des informations ou interagit directement avec l'utilisateur inaverti pour obtenir des identifiants ou autre en fonction de l'application à laquelle l'utilisateur accorde sa confiance.

7. https://en.wikipedia.org/wiki/Code_injection

3.6 Ressources, solutions et outils existants

3.6.1 Le travail d'OWASP

La fondation [OWASP](#) (*The Open Web Application Security Project*) est une organisation à l'échelle mondiale qui aspire à améliorer la sécurité informatique des systèmes et logiciels. L'IoT étant en plein essor, un projet a vu le jour sur la plateforme, décrit comme « conçu pour aider les fabricants, développeurs et utilisateurs à mieux comprendre les problèmes sécuritaires liés à l'IoT, et pour permettre aux utilisateurs de prendre de meilleures décisions relatives à la sécurité quand ils déplient et travaillent avec des technologies de l'IoT ». Ainsi, plusieurs sous-projets plus spécifiques sont destinés à peupler le projet principal, certains étant encore actuellement en développement.

L'équipe travaillant sur le projet a établi [une liste](#) des principes relatif à la sécurité dans l'IoT. Plusieurs guides ont également été écrits, orientés chacun selon un point de vue différent : pour le fabricant, le développeur et l'utilisateur. D'autres ressources écrites sont disponibles, mais encore aucun outil développé autour de la sécurité n'a vu le jour comme c'est le cas pour d'autres domaines (les applications web particulièrement puisque originellement l'organisation ciblait ce secteur).

3.6.2 La standardisation de protocoles adaptés sécurisés

La sous-section 1.3 présentait les protocoles utilisables dans l'IoT, adaptés aux restrictions des équipements qui ne permettent pas toujours une compatibilité avec la *pile TCP/IP* classique. Lors de la conception de ces protocoles, la composante sécurité a été prise en compte et des mécanismes y ont été introduits pour l'assurer [21] (ou alors des chercheurs ont fait ce travail par la suite).

6LowPAN propose déjà des mécanismes utiles pour garantir la sécurité [1]. Pour éviter les attaques par fragmentation, les options d'en-têtes *Timestamp* et *Nonce* sont un premier rempart. La découverte des voisins d'un noeud est effectuée de façon sécurisée grâce à LSEND (*Lightweight Secure Neighbor Discovery protocol*), version allégée de SEND. Des chercheurs ont également élaboré une version allégée de la suite de protocoles IPsec afin de garantir des communications sécurisées d'un noeud à l'autre dans un réseau 6LowPAN [21]. D'autres ont proposé un système d'établissement de connexion sécurisée à partir d'un échange de clés [21].

RPL, utilisé pour le routage dans les réseaux de capteurs notamment au dessus de 6LowPAN, dispose également de mécanismes de protection. Il s'agit dans ce cas de garantir que quand un noeud sélectionne son parent pour rejoindre la topologie, ce dernier soit un noeud légitime (éviter le *spoofing*). Des solutions d'autres chercheurs sont présentées par Huichen [21] afin de sécuriser la génération des tables de routage dans RPL. Un exemple est un schéma de topologie authentifiée appelée TRAIL (*Trust Anchor Interconnection Loop*) qui permet la découverte et l'isolation des noeuds falsifiés et où chaque noeud peut, en envoyant un message particulier, valider son chemin jusqu'à la racine.

CoAP, tout comme son homologue HTTP, peut s'appuyer sur une couche intermédiaire plus basse pour assurer un transport sécurisé (chiffré pour préserver la confidentialité). Il s'agit de DTLS (*Datagram Transport Layer Security*), un équivalent de TLS plus adapté aux communications d'équipements restreint. La principale différence réside dans le fait que TLS soit basé sur TCP, au contraire de DTLS qui l'est sur UDP, toujours dans l'optique d'utiliser le minimum de ressources. CoAP offre le choix entre quatre modes relatifs à la sécurité : *NoSec*, *PreSharedKey*, *RawPublicKey* et *Certificate*.

3.6.3 La détection des équipements en réseau et de leurs vulnérabilités

Face à un réseau réel, disons local à une habitation possédant des équipements connectés (c-à-d orientée smarthome), il est difficile de se faire une idée de tous les objets y étant connectés et de quelles sont leurs fonctionnalités. Des outils sont nécessaires afin de découvrir le réseau et établir quels équipements peuvent potentiellement être la cible d'attaques. Il en existe une flopée ([sectools.org](#) en maintient un classement fourni), libres ou commerciaux, possédant une interface graphique ou non, fonctionnels sur des OS variés. Outre ces logiciels complets au champ d'action assez étendu, des outils de scanning plus spécifiques à des cibles (routeurs, serveurs, ...) et services (Web, SQL, mail, ...) sont trouvables dans des dépôts dédiés sous la forme de scripts ([Scanners-box](#) est dédié à leur référencement). Des exemples sont donnés ci-dessous, classés en deux catégories [20] en fonction de la manière dont le scan est effectué. Une attention particulière est portée aux outils dédiés à la détection des failles sur les équipements IoT d'un réseau s'il en existe qui soient dédiés.

3.6.3.1 Le scan passif et la prise d'empreintes

L'objectif est d'en apprendre un maximum sur les équipements connectés au réseau sans les interroger directement. Cela se fait en écoutant toutes les transmissions passant sur un média utilisé par ces équipements, les capturant, les stockant et les agrégeant. Cela est donc assimilable à la pratique du sniffing. L'utilisation qui est faite de ces données varie en fonction de l'application : soit elles sont étudiées telles quelles car c'est le trafic qui est intéressant ([analyseurs de trafic](#)), soit elles sont utilisées pour faire de la « prise d'empreinte » ([TCP/IP stack, OS fingerprinting](#)) afin de caractériser les nœuds du réseau, soit les deux. Il peut également y avoir, suite à la phase d'écoute passive, des actions entreprises (par exemple si le processus a déterminé qu'un hôte présente une faiblesse exploitable).

Un analyseur de paquet populaire est Wireshark, qui utilise l'API *pcap* pour capturer les paquets transitant par le réseau sur lequel une interface désignée est placée en mode promiscuité (c-à-d fait remonter au système tous les paquets qui lui parviennent). Beaucoup d'autres utilitaires de capture (Kismet, tcpdump, ngrep, etc.) utilisent l'API *pcap*, implémentée dans la librairie *libcap* (systèmes UNIX) ou WinPcap (port de libcap vers Windows), il s'agit donc d'une ressource importante. Wireshark permet un filtrage assez fin des lots de paquets capturés (plus d'un millier de protocoles reconnus), permettant une grande flexibilité en fonction du but de la capture (debug, recherche d'une information, surveillance du trafic, etc.).

Les logiciels destinés à la capture d'empreintes et uniquement à cela sont assez peu nombreux (la plupart intègrent d'autres fonctionnalités actives). [p0f](#) est l'outil minimaliste le plus efficace. Il utilise diverses métriques extraites des en-têtes IPv4/6, des en-têtes TCP, des dynamiques observables dans le 3-way handshake de TCP et des données exploitables au niveau de la couche application (HTTP, SMTP, FTP, etc.). Combinées à une base de données de signatures récupérées expérimentalement, ces mesures permettent de déduire quels systèmes opèrent derrière les nœuds du réseau communiquant (là où un outil comme Nmap peut en être incapable). Outre cela, p0f est capable de déduire d'autres informations :

- présence d'un pare-feu, d'un système de NAT ou *load balancer*, de proxy applicatifs
- mesures du temps d'activité d'un système, sa distance (grâce au TTL), ses préférences au niveau de la langue, etc.
- les services utilisés par les clients et serveurs qui déclarent des champs comme X-Mailer ou User-Agent

D'autres outils existent, leurs fonctionnalités s'étendant au-delà de ce qui est considérable comme outil passif. On peut citer notamment :

- [Ettercap](#) utilise la capture de paquets pour intercepter des informations sensibles, et d'autres opérations avec pour but final d'effectuer une attaque de type *Man-In-The-Middle* (MITM)
- [Scapy](#) et [Kamene](#) (fork du projet), des utilitaires en Python, permettent la capture et le reforgeage de paquets, la prise d'empreintes, le lancement d'attaques, etc. (le tout de façon modulable)

3.6.3.2 Le scan actif

Ces logiciels tentent d'établir des connexions avec les hôtes à analyser, à la recherche d'un service répondant donc potentiellement ouvert, non protégé et dont il est possible de soutirer des informations. Une liste de numéros de ports pertinents est construite par le scanneur, utilisée par ce dernier lors de l'analyse des hôtes qu'il trouve dans le réseau. Cette technique basique est appelée **balayage de port**. Se servant de ces scanneurs comme appoint, les **scanneurs de vulnérabilités** vont plus loin en tentant de découvrir pour chaque hôte analysé les failles de sécurité auxquelles ils sont exposés (au niveau d'une application, du système d'exploitation ou des interfaces réseaux).

L'outil **Nmap** est dédié au balayage de ports et à l'acquisition directe d'informations (version, composants matériels utilisés, etc.) concernant le système et les services ouverts d'un hôte analysé. La détection de vulnérabilités est rendue possible grâce aux interactions scriptées avec la cible en utilisant le Nmap Scripting Engine (NSE), un exemple de module Nmap y étant dédié est **Vulscan.nse**. Nmap a d'autres fonctionnalités pratiques telles qu'une interface graphique complète (**Zenmap**) et la personnalisation des formats de sorties pour les résultats (filtrage, langages balisés, etc.).

OpenVAS est une branche du projet **Nessus**, ce dernier étant passé sous licence commerciale. Ce sont les frameworks pour la détection de vulnérabilités sur un réseau les plus populaires. Tenable® (qui possède Nessus) propose des **services de scan dédiés aux équipements IoT** utilisés dans l'industrie, mais OpenVAS n'offre rien de similaire.

Le cas de **Lynis** est un peu particulier. Effectivement, en opposition aux scanneurs en réseau comme OpenVAS, il s'agit d'un scanneur de vulnérabilités à lancer sur l'hôte à analyser. Cela permet une inspection beaucoup plus en profondeur des services et des vulnérabilités qu'il présente. Il tourne sur presque tous les systèmes UNIX et même sur un système limité tel qu'une carte Raspberry Pi ou un équipement relatif à l'IoT qui est capable de supporter le même type de charge. Lynis peut donc être un outil intéressant si on possède le contrôle de l'équipement à analyser (légitimement ou non).

Dédié à la détection des vulnérabilités des équipements IoT, **IoTSeeker** est un outil libre et écrit en Perl, supporté par les systèmes Linux/Mac OS. Il est axé sur les vulnérabilités d'authentification, la découverte de mots de passe laissés par défaut par les utilisateurs imprudents (sur base des listes des identifiants par constructeur évoquée dans la sous-section 3.1).

RouterSploit est un framework dédié à la détection et l'exploitation de failles dans les systèmes embarqués. Comme son nom l'indique, la cible est davantage les routeurs que les équipements de l'IoT, cependant quelques modules ciblent notamment les caméras et d'autres équipements connectés insolites (projecteur, laveur de laboratoire). L'outil de pénétration **Nettacker** développé par OWASP, dont le but est le soutirage et l'agrégation d'informations d'une cible pour en dresser une liste des failles de sécurité, promet également la sortie future d'un module spécifique à l'IoT.

4 Développement d'une application *IoTMonitor* - introduction et concepts

4.1 Présentation générale

4.1.1 Contexte et environnement ciblé

Ce travail englobe la thématique de la sécurité des équipements en réseau dans l'environnement domestique, en mettant l'accent sur les différents types de pare-feux. Effectivement, ces derniers peuvent palier en partie aux faiblesses récurrentes des équipements IoT qui peuplent les réseaux domestiques. L'aboutissement final du projet est de développer un outil qui puisse contribuer à la protection de ces réseaux. Cependant, l'analyse de trafic avancée (au delà du simple pare-feu à état) telle qu'utilisée dans les solutions commerciales est une technologie non triviale qui sort des limites de ce projet (théorie de *machine learning*, construction de base de données de signatures, etc.). À l'inverse, se limiter à un simple pare-feu à état ciblant un protocole de l'IoT peut sembler réducteur vis-à-vis de toutes les menaces planant sur ces équipements qu'un tel pare-feu ne pourrait contrer.

La sous-section 3.6 met en évidence quelques outils qui peuvent aider à analyser le réseau domestique et détecter les failles exploitables des équipements qui le peuplent. Ces *scanners*, conjugués à des systèmes de prévention d'intrusion *NIDS*, permettraient d'améliorer grandement la sécurité du réseau domestique, tout en n'effectuant aucune tâche de filtrage direct du trafic (c-à-d n'étant pas bloquant comme le serait un véritable pare-feu). En découle l'idée motrice de l'application à développer : prodiguer un outil permettant **l'analyse, la surveillance et l'agrégation des caractéristiques des équipements** d'un réseau domestique, basé sur des **scanneurs** existants ciblant **différents aspects** de sa sécurité. Le système doit donc être conçu de façon modulaire, de sorte que **l'intégration d'un nouveau scanneur** ne requiert comme travail que la traduction (parsing) de la sortie dudit programme vers des actions et éléments interprétables par l'application.

Par exemple, l'application pourrait intégrer des scanneurs tels que *nmap*, *IoTSeeker*, etc. qui interrogent directement les équipements (actifs) et des programmes de prise d'empreintes/analyse de trafic tels que *Snort*, *p0f*, etc. (passifs). Chacun de ces programmes fournit un feedback sur sa sortie, que ce soit une fois son exécution terminée ou de façon continue jusqu'à interruption de l'exécution. Une fois parsées, les informations obtenues peuvent servir à alimenter une représentation virtuelle du réseau maintenue par l'application (renseignant les caractéristiques des équipements qui le peuplent) ou constituer une alerte à faire remonter jusqu'à l'utilisateur si une faille ou menace est détectée. Le framework d'intégration définit comment ajouter ces programmes à l'environnement de l'application, prodiguant une surcouche telle qu'il n'est nécessaire de fournir que la **commande** liée au programme, la façon de les lui passer et quels **paramètres** du programme utilisé et la **méthode de traitement** de la sortie.

4.1.2 Objectifs de l'application

4.1.2.1 La surveillance du réseau domestique

L'application doit pouvoir être utilisée telle quelle depuis un équipement connecté au réseau domestique (similaire ou supérieur en terme de ressources à une carte Raspberry PI), avec un lot de programmes (sous entendu scanneurs dans la plupart des cas) déjà intégrés et prêts à l'emploi par une personne non initiée. Les interactions se font principalement sur deux composants décris plus en détail dans la sous-section 4.2 : la routine de surveillance automatisée et la carte du réseau. En terme de concepts :

- **La routine de surveillance** est l'ensemble des scanneurs qui doivent être utilisés pour analysé le réseau quand l'application est lancée. Elle se divise en deux sous-ensembles, l'un destiné aux programmes tournant en continu (analogie avec les scanneurs passifs, la prise d'empreintes) et l'autre aux programmes/scripts qui interagissent directement avec les équipements du réseau pour en sortir un feedback une fois leur analyse terminée (analogie avec les scanneurs actifs). L'automatisation pour le premier se fait simplement en associant un état commutable *en cours/arrêté* à chaque programme qui contrôle le lancement/interruption du processus sous-jacent. Pour le second, il s'agit d'un timer d'expiration associé à chaque programme dans l'ensemble, au terme duquel le programme est exécuté (se termine et fournit un feedback sur sa sortie) et son timer restauré à sa valeur initiale (on peut donc voir cela comme une file de programmes à exécutés les uns à la suite des autres). Ces deux ensembles sont manipulables, des programmes peuvent être ajoutés/retirés et leur état modifié en agissant sur ce qui sera désigné par la **Routine** qui regroupe ces deux ensembles **Panel** et **Queue** respectivement.

- **La carte du réseau** est le résultat de l'agrégation des informations parsées sur les sorties des programmes de la Routine (donc automatiques) et des entrées manuelles de l'utilisateur. Il s'agit de résumer au mieux d'une part les informations récoltées concernant les équipements (adresses MAC, IP, hostname, ports ouverts, etc.) et d'autre part les alertes associées à un équipement particulier (CVE existante pour le modèle, service utilisé non sécurisé ou sujet à une faille connue, etc.). Cette carte virtuelle du réseau a pour objectif de fournir une vue d'ensemble du réseau domestique (limitée au LAN) pour l'utilisateur, afin qu'il puisse y exercer un contrôle plus pertinent.

4.1.2.2 La conception d'un framework d'intégration

Puisque l'application fonctionne uniquement en se basant sur des programmes extérieurs arbitraires, il est nécessaire de définir un cadre assez rigoureux pour les intégrer au système. Cela se fait via une surcouche d'abstraction désignée par le terme **Module**. Deux archéotypes sont envisagés pour un Module, en fonction de la nature de l'exécution du programme à abstraire : soit celle-ci est continue et produit un feedback régulier jusqu'à terminaison du processus, soit il s'agit d'une exécution unique destinée à se terminer avec un feedback des résultats e, sortie. Par analogie avec les types de scanneurs, on désigne l'archéotype de Module à suivre correspondant par **Module Actif** et **Module Passif**. La surcouche correspond dans les faits à hériter d'une classe abstraite dont plusieurs méthodes doivent être implémentées, cette partie pratique est approfondie dans la sous-section 4.3.

De nouveaux programmes peuvent ainsi être intégrés et utilisables dans l'application, à condition d'écrire une classe héritant de la classe correspondant soit aux modules actifs, soit aux modules passifs en fonction de la nature du programme. Dans ces classes supérieures, la façon dont l'exécution du programme sous-jacent doit être traitée est déjà encadrée (bien qu'il soit aussi possible de réécrire cela). Le schéma d'exécution est le suivant :

- **Module actif** : un thread est lancé pour dissocier le fil d'exécution général de l'application. Dans ce thread, un processus fils exécutant le programme est instancié, avec sa sortie récupérable. Une fois l'exécution terminée, le thread se termine en appelant une fonction de callback qui est celle définie pour le parsing du programme.
- **Module passif** : le travail est divisé en deux parties, la première se charge de lancer un processus fils exécutant le programme dont la sortie est redirigée vers un pipe et la seconde est chargée de lire le contenu de ce pipe et d'appeler la fonction de parsing dessus à intervalle de temps régulier.

Comme énoncé lors de son introduction, la Routine est composée de deux ensembles de modules : Panel et Queue. Le premier est donc destiné à accueillir des modules respectant l'archéotype **passif**, tandis ce que le second s'attend à des modules présentant l'archéotype **actif**. Sur base de cette hypothèse, un contrôle cohérent de leur exécution peut être établi : par exemple l'action de *mettre en pause* un module passif dans le Panel et actif dans la Queue a une sémantique différente. Mettre en pause un module passif équivaut à envoyer un signal d'interruption au processus sous-jacent tournant en continu en arrière-plan. Mettre en pause un module actif dans la Queue équivaut à ne plus décrémenter son timer d'expiration associé, au terme duquel il était censé lancer une exécution du programme qu'il abstrait.

L'application référence tous les modules disponibles dans une **Librairie** de modules. Il ne s'agit, en arrière-plan, que d'un simple fichier XML décrivant tous les modules qui doivent être chargés dans un format strict. Afin d'intégrer de façon constante un module pour qu'il soit utilisable dans l'application, deux options sont possibles : soit écrire manuellement la description du module et l'ajouter au fichier correspondant (plusieurs librairies de modules peuvent être envisagées), soit simplement appeler une méthode dédiée sur une instance du module qui en créera le descriptif automatiquement. Si le module a été correctement écrit en suivant le strict minimum imposé par le framework, le descriptif du module sera ajouté à la librairie courante et le module facilement manipulable dans l'application (pour par exemple l'ajouter à la Routine de surveillance).

4.2 Composants du système de surveillance

Afin de mieux visualiser les différents éléments qui composent le cœur de l'application, une maquette d'interface graphique a été élaborée. Celle-ci représente ce à quoi devrait correspondre l'écran de gestion principal de l'application. Effectivement, afin de pouvoir être utilisée dans un contexte restreint (tel que sur une carte Raspberry PI), l'application a d'abord été dotée d'une interface en ligne de commande en priorité (voir point 5.2.5). Une partie des composants principaux sont repris dans la figure 9.

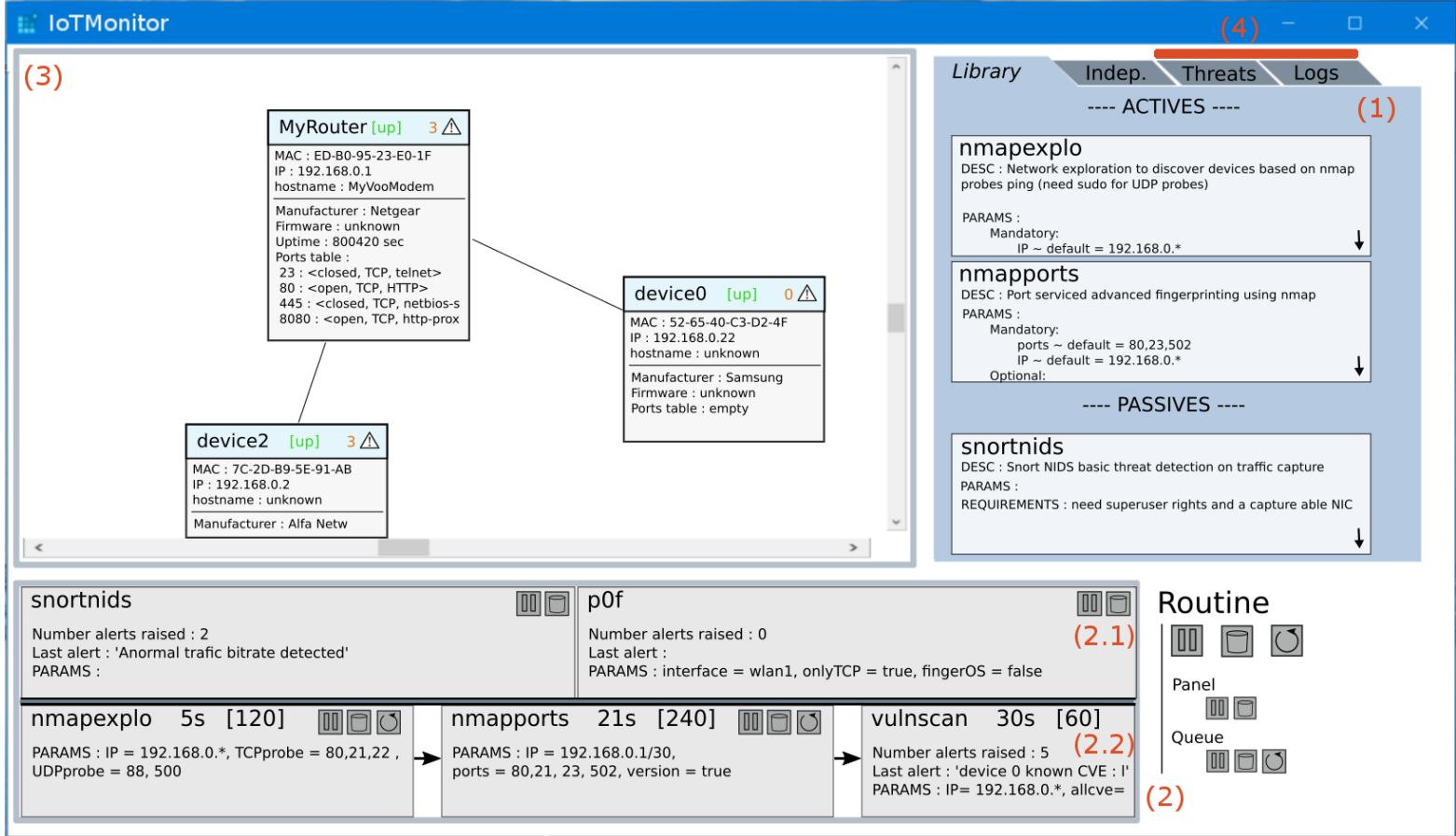


FIGURE 9 – Maquette GUI de l'application en reprenant les principaux éléments

4.2.1 Modules actifs et passifs - abstraction des programmes

Les Modules constituent la couche d'abstraction pour les programmes sous-jacents et sont régis par un framework spécifique à l'archéotype actif ou passif. Par programme il est entendu : *un programme ou script installé et fonctionnel sur le système, exécutable depuis un terminal et dont la sortie est récupérable*. Plusieurs modules distincts peuvent abstraire le même programme, mais dont les utilisations spécifiques sont différentes. Par exemple, le programme *nmap* est accompagné du NSE (Nmap Script Engine), un framework permettant d'intégrer des scripts variés à l'exécution d'un scan nmap. Nmap peut donc être utilisé en combinaison avec différents scripts dont les objectifs (et donc les paramètres et sorties du programme) sont très hétérogènes. Cela implique que des couches d'abstraction distinctes devraient être écrites afin de les représenter par différents modules dans l'application.

Les modules sont manipulables comme des unités dans l'application. Ils sont référencés dans la Library qui est le point d'entrée à partir duquel les instancier dans l'application. Elle est représentée dans la maquette par le label (1). À partir de là, un module peut soit être inséré dans la routine de surveillance si on souhaite son utilisation régulière automatisée, ou dans un ensemble à part, dit indépendant, où il sera exécuté directement une unique fois (ce qui a plus de sens pour un module actif, mais reste possible pour un module passif).

Les attributs qui devraient être définis lors de l'écriture d'un module doivent au minimum comporter :

- Un id unique parmi tous les modules déjà référencés dans la Library
- Une description textuelle du rôle du module
- Une indication de l'archétype associé au module : à considérer comme actif ou passif
- La commande qui permet d'appeler l'exécutable sur le système
- Une définition des paramètres ajoutables à la commande, avec pour chacun :
 - un indicateur booléen indiquant s'il est indispensable (toujours rajouté à la commande)
 - une valeur par défaut, qui sera utilisée si le paramètre est indispensable mais non fourni
 - le flag associé à l'option du programme en ligne de commande si existant
- Additionnellement, une description textuelle de chaque paramètre
- La méthode de parsing

Le champ d'action d'un module tel qu'il devrait être suivant le framework est le suivant :

1. Appeler le programme avec les paramètres qui lui ont été communiqués ou ceux définis comme par défaut
2. Dissocier le fil d'exécution en exécutant le programme sous-jacent dans un nouveau thread
3. Maintenir une référence vers ce processus de sorte à pouvoir communiquer avec (signal d'interruption, etc.)
4. Récupérer les informations que le programme renvoie sur sa sortie standard et ses erreurs
5. Parser cette sortie et la traduire vers l'application afin d'y intégrer l'information :
 1. **Logging** : informations diverses à inclure dans les logs de l'application (debug, ...)
 2. **Alertes** : menace détectée par le programme, à ne pas simplement logger (associer à un équipement particulier, création et envoi d'un mail au propriétaire, ...)
 3. **Modifications** : un module a l'accès en lecture et écriture sur la carte du réseau et ses éléments, permettant leur mise à jour automatique depuis les sorties des programmes

Pour plus de détails concernant l'implémentation concrète du système de modules, se référer à la sous-section 4.3.

4.2.2 Routine - automatisation de la surveillance

La Routine est l'ensemble des modules sélectionnés pour surveiller le réseau domestique de façon automatisée, en répétition continue. On la retrouve associée dans la maquette au label (2). La Routine courante est remplie par l'utilisateur depuis la Library et peut être paramétrée, modifiée et contrôlée par celui-ci à volonté. Elle maintient deux sous-ensembles vers lesquels les modules sont aiguillés en fonction de leur archétype : le Panel des passifs (2.1) et la Queue des actifs (2.2). À l'ajout d'un module, un index unique parmi le sous-ensemble est associé (par défaut l'id du module) et toute interaction avec le module peut se faire via ce **setid** (respectivement un **pid** et un **qid** dans les deux sous-ensembles). Un même module peut être présent en de multiples exemplaires, tant qu'il s'agit d'instances différentes. Cela permet, par exemple, d'avoir deux instances du même module mais qui considéreront des paramètres différents pour le programme sous-jacent.

Le Panel maintient les modules respectant l'archétype *passif*, c'est-à-dire dont le programme sous-jacent est destiné à être exécuté de façon continue jusqu'à interruption du processus. En plus d'un **pid**, à chaque module est donc associé un état "en cours" ou "pause", et au Panel également (panel "en cours" \Leftrightarrow au moins un de ses modules est "en cours"). Cet état est commutable pour chacun de ces éléments, le passage à l'état "en cours" pour le Panel revient à lancer l'exécution de tous ses modules (idem pour "pause").

La Queue contient les modules dits *actifs*, dont le programme sous-jacent effectue un travail en temps fini au terme duquel il a terminé son exécution. Le but de la Queue est d'organiser la séquence des modules qui seront lancés. Pour cela, en plus du **qid** attribué à chacun de ses modules, un timer d'expiration est initialisé (à une valeur fournie par le module et modifiable). Il est décrémenté chaque seconde jusqu'à valoir 0, alors le module est lancé (entendu l'exécution du programme sous-jacent), et son timer est réinitialisé. L'ensemble des modules de la Queue peut alors être visualisé comme une file, où les modules sont agencés selon leur temps restant. On visualise ainsi la séquence d'exécution prochaine, totalement automatisée.

4.2.3 Netmap - recensement des équipements du réseau

La carte du réseau est l'élément vers lequel doivent converger toutes les informations concernant le réseau, obtenues grâce aux programmes sous-jacents ou entrées par l'utilisateur lui-même. Elle a pour objectif de les agréger pour produire une représentation interne du réseau domestique cohérente, comme représenté dans la maquette par (3). Les équipements physiques y sont représentés par des **Instances Virtuelles** (IV), chacune destinée à correspondre à un unique équipement et regroupant un maximum des informations le concernant. Cet ensemble d'IVs est maintenu dans la **Netmap**, où elles sont indexées par un identifiant `mapid` attribué automatiquement `device0`, `device1`, ... modifiable par l'utilisateur s'il le désire en fonction de ce qu'est réellement l'équipement pour lui.

Les informations stockées dans les IVs sont appelées des **champs**. Ceux-ci peuvent être créés et remplis de façon automatisée par les modules (c'est un travail de la fonction de parsing des sorties à définir pour chaque module), mais également par l'utilisateur lui-même s'il souhaite étayer les informations concernant ses équipements. Quelques champs sont plus importants que d'autres, car plus pertinents dans un processus d'identification de l'équipement concerné par un paquet d'informations. Par exemple, un scanner de ports des équipements du réseau envoie sur sa sortie l'IP de chaque équipement analysé et une table des ports ouverts. Pour associer cette information à chaque IV concernée, il va falloir identifier dans la Netmap l'IV unique qui renseigne déjà cette IP. S'il n'y en a pas, une nouvelle IV est créée pour contenir l'information. Il est possible que l'équipement réel était en effet déjà référencé dans la Netmap, mais son champ IP vide. Il s'agit donc d'un problème non trivial, une première étape pour le résoudre est de définir dans quel ordre certains champs sont pertinents pour évaluer une correspondance.

Les champs principaux sont les suivants (par ordre de pertinence) :

1. Adresse MAC (unique pour chaque équipement, sauf en cas de spoofing)
2. Adresse IP (allouée généralement dynamiquement)
3. Nom d'hôte `hostname` (aucune garantie d'unicité, mais assez probable)

D'autres champs, considérés comme divers et auxquels de nouveaux peuvent venir se greffer à volonté :

- Fabriquant
- Modèle
- Firmware
- ...

Afin de limiter ce besoin d'identifier l'instance virtuelle relative à une information obtenue du réseau, une mécanique simple est prévue. Quand un module est lancé, il est possible de lui fournir en paramètre quelles sont les IVs auxquelles l'exécution est relative. Cela permet deux améliorations : accéder à des informations sur les équipements déjà connues pour paramétriser l'exécution du programme et à interpréter sa sortie en fonction d'IVs déjà connues (donc de ne pas avoir à chercher une correspondance). Cet aspect est à gérer dans l'implémentation de chaque module.

À chaque instance virtuelle est également associée une Table des ports. Celle-ci maintient de l'information pour chaque port qui a été détecté. L'indexation se fait simplement par le numéro associé au port, et les valeurs principales à y associer sont le nom du service correspondant, le protocole de transport utilisé et son état (`open`, `closed`, ...). Il y a également la possibilité d'y associer des informations diverses à volonté telles que la version du service, le taux de confiance en cette information, etc.

Une entrée dans cette table est donc de la forme :

`numPort : <service, protocole, état>, divers {(serv_version, val), (conf, val), ...}`

4.2.4 Centre d'événements - logging et lancement d'alertes

On distingue deux catégories d'informations qui doivent être enregistrées par l'application : les logs purs et les événements représentés par un objet **Event**. Les logs sont de simples informations à écrire dans un fichier ou une sortie donnée, tels que des du debug, des codes de retour à la terminaison des programmes et autres informations lors de leur exécution, etc. Les Events sont quant à eux capturés par l'application et instanciés pour être manipulés par la suite. Les logs interviennent principalement dans le cadre de l'exécution des programmes sous-jacents aux modules, et sont donc être utilisables depuis l'endroit de leur définition. La place de ces éléments dans la maquette est indiquée par (4).

Les Events peuvent être de deux natures, et constituent un moyen pour les modules de communiquer les modifications qu'ils effectuent sur les éléments de l'application. Soit ce sont des événements de modification, soit de menace. Le premier cas fait référence aux modifications sur la Netmap et ses instances virtuelles, notamment la création d'une IV ou un ajout ou changement de valeur d'un de ses champs. L'objet Event contient principalement des informations telles que la valeur précédente et la nouvelle, le module qui a fait la modification. Un Event de menace est utilisé pour décrire une détection de menace au niveau de la sécurité du réseau domestique. Elle maintient des informations sur la nature de la menace, le possible patch à appliquer, le niveau de danger, etc. qui doivent être remplies par le module qui lance l'alerte.

L'élément de centralisation des Events de l'application est appelé **EventCenter**. Il est accessible par un simple appel depuis n'importe où dans l'application et permet de créer des Events qui y seront enregistrés. Il est alors possible d'appliquer des fonctions de recherche et filtrage sur ces événements, par exemple pour trouver toutes les modifications sur une instance virtuelle donnée. Un nombre fixe limite le maximum d'événements gardés en mémoire, les plus anciens laissant leur place aux plus récents.

Pour éviter la perte d'événements importants concernant des équipements du réseau de la sorte, il est également possible d'accéder à l'EventCenter en passant par la Netmap. Un Event créé par ce biais, s'il concerne une IV de la Netmap, sera conservé de façon permanente dans la Netmap en plus de la journalisation temporaire dans l'EventCenter. L'indexation des Events dans la Netmap se fait simplement par le même index d'IV (son **mapid**). L'utilisateur peut ainsi facilement visualiser les modifications et menaces impliquant directement ses équipements.

4.3 Modèle d'abstraction des logiciels

4.3.1 Modèle général d'abstraction

Bien que l'on considère deux archétypes différents, les modules actifs et passifs doivent implémenter plusieurs méthodes communes. On définit donc ce qui équivaut à une interface générale dont deux classes abstraites héritent. Le framework pour l'intégration des modules est défini de cette façon. La figure 10 reprend ce framework, illustrant quelles méthodes doivent être implémentées pour que l'application puisse utiliser le nouveau module écrit.

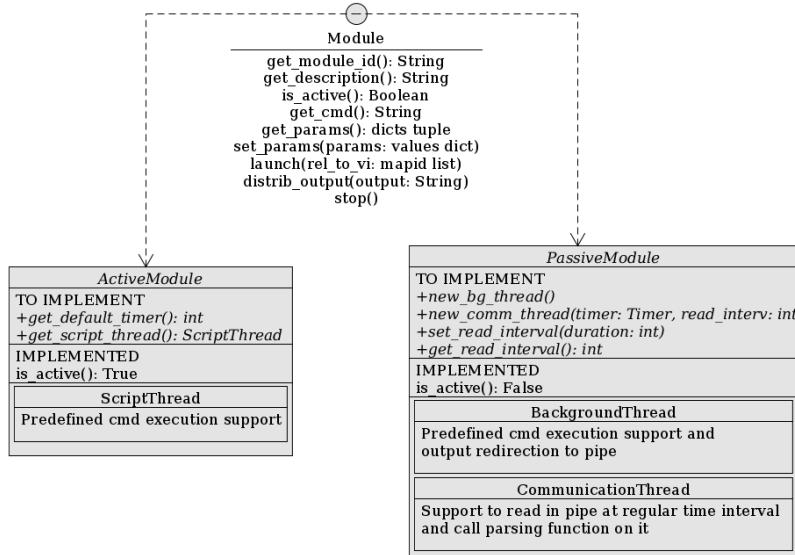


FIGURE 10 – Structure orientée objet du modèle d'abstraction

À noter que chaque module devrait garder une référence vers les threads dans lesquels il a lancé une exécution de son programme sous-jacent (plus de détails aux points suivants). Les méthodes remarquables de l'interface *Module* ont la sémantique suivante :

- *get_cmd()* : retourne l'alias qui sert à appeler le programme en CLI (comme lors d'une utilisation normale)
- *get_params()* : obtention de 3 éléments : paramètres actuels, par défaut et leur description textuelle
- *set_params()* : ajustement de quels paramètres et de leur valeur à fournir à une exécution du programme
- *launch()* : lance une exécution du programme dans un nouveau thread, avec les paramètres courants du module. Les définitions pour un schéma d'exécution classique sont déjà implémentées dans les classes abstraites de l'archéotype correspondant (ce sont des objets instanciables avec certains paramètres).
- *distrib_output()* : méthode de parsing appelée sur la sortie de l'exécution d'un programme, effectue le travail de traduction vers les éléments de l'application
- *stop()* : interrompt l'exécution de tous les programmes sous-jacents dans les threads liés à l'instance du module

4.3.2 Design pour un module actif

Le schéma d'exécution classique pour l'archéotype actif est simple : l'appel à *launch()* doit lancer l'exécution du programme sous-jacent avec les paramètres actuels du module et dans un nouveau thread, auquel une fonction de callback (*distrib_output()*) est passée en paramètre. À la fin de l'exécution, la sortie du programme est récupérée et l'appel à la fonction de callback est effectué sur cette sortie (donc le parsing se fait dans la fin du fil d'exécution du thread). Ce comportement peut être obtenu grâce au support déjà implémenté dans la classe abstraite *ActiveModule* dont un module concret doit hériter. Il reste cependant des méthodes spécifiques au module qui doivent y être implémentées (notamment celles de l'interface *Module*). La figure 11 résume cette information.

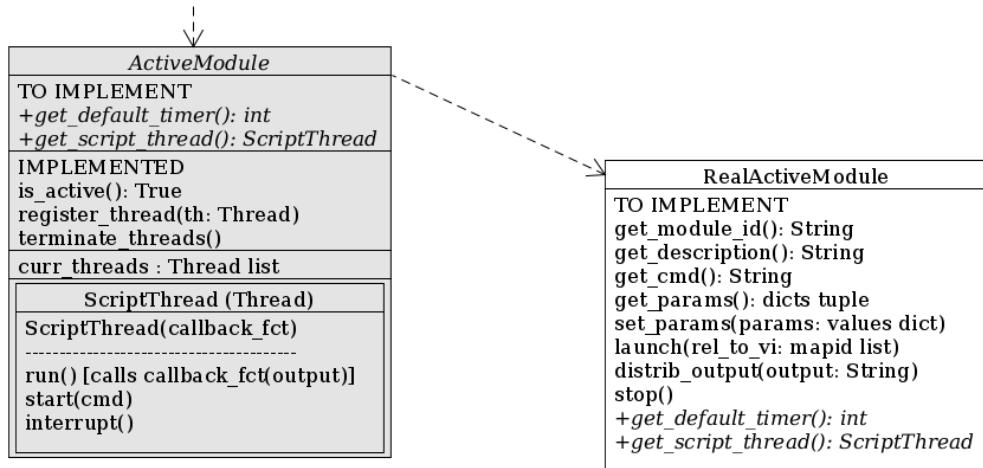


FIGURE 11 – Framework d'écriture d'un module actif

Le code d'un module actif devrait être structuré de la sorte (si utiliser les facilités déjà implémentées est désiré) :

1. *get_script_thread()* doit retourner une instance de *ScriptThread* paramétrisée avec la fonction *distrib_output()* du module.
2. *launch()* construit la commande d'appel au programme *cmd* depuis les paramètres actuels du module, telle qu'elle serait invoquée en ligne de commande (en utilisant *get_cmd()* + les paramètres actuels dans *get_params()*).
3. *launch()* instancie le conteneur de l'exécution avec *get_script_thread()* et invoque sa méthode *start()* avec *cmd*, ce qui va lancer un nouveau thread dans lequel le processus du programme sous-jacent est créé. À la fin de son travail (*run()* est bloquant dessus), la méthode *distrib_output()* est appelée sur sa sortie.
4. *launch()*, après avoir créé et dissocié le thread du fil d'exécution, l'enregistre dans l'ensemble des exécutions courantes du module avec *register_thread()*.

Additionnellement, la méthode *get_default_timer()* renseigne la valeur par défaut que devrait prendre le timer d'expiration d'une instance du module placée dans la Queue.

4.3.3 Design pour un module passif

Pour un programme correspondant à l'archétype passif, le schéma d'exécution est complexifié car il faut pouvoir récupérer les sorties du programme au fur et à mesure qu'elles sont produites et de façon régulière. Effectivement, on ne peut pas attendre la terminaison du programme qui marque le fait qu'il y a quelque chose à lire sur sa sortie puisqu'il n'est pas censé terminer son exécution de lui-même. La solution choisie est de séparer le schéma d'exécution en deux parties. L'une contrôle le processus sous-jacent et redirige sa sortie vers un pipe temporaire pendant que l'autre partie y lit les sorties à intervalle régulier et les parse. Comme pour l'archétype actif, des facilités sont fournies pour gérer ce travail, illustrées par la figure 12.

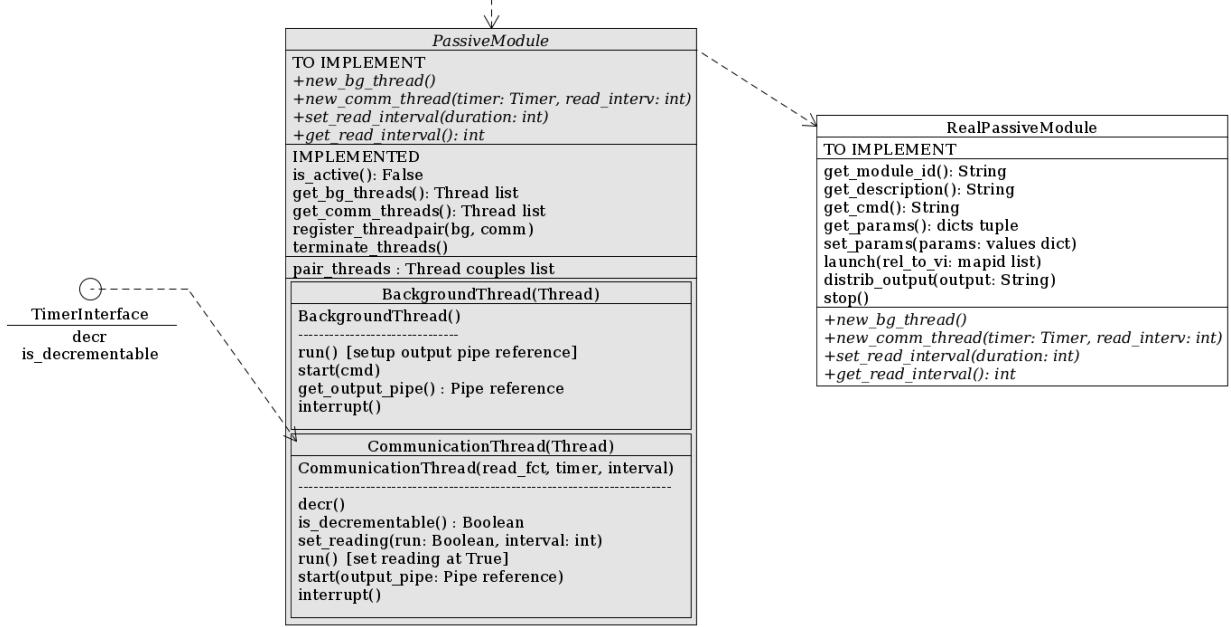


FIGURE 12 – Framework d'écriture d'un module passif

Pour utiliser correctement ces facilités, l'implémentation du module devrait respecter ces recommandations :

1. `new_bg_thread()` et `new_com_thread()` retournent respectivement une instance de `BackgroundThread` et `CommunicationThread`. Cette dernière doit être paramétrisée avec un objet `Timer`, l'intervalle de temps auquel la lecture doit avoir lieu et la fonction de parsing `distrib_output()` appelée sur la sortie lue
2. `set_read_interval()` et `get_read_interval()` sont relatifs à l'intervalle entre deux lectures qui doit être > 0
3. Comme pour les modules actifs, `launch()` devrait construire la commande `cmd` telle qu'utilisable en CLI
4. Dans un premier temps, `launch()` appelle la méthode `start()` d'une instance de `BackgroundThread` en lui passant `cmd` pour démarrer le processus du programme sous-jacent, tout en redirigeant sa sortie vers un pipe alors créé à cette fin (`output_pipe`). Une référence vers ce dernier est maintenue.
5. En second lieu, `launch()` devrait attendre que le processus ait terminé son initialisation, c'est-à-dire que `output_pipe` soit créé et accessible en lecture. Alors le `CommunicationThread` peut être instancié en le lui passant en paramètre.
6. Finalement, `launch()` devrait enregistrer cette paire d'objets Thread dans la liste prévue à cet effet avec un appel à `register_thread_pair()`

À noter qu'ici, la méthode de `BackgroundThread run()` lançant le sous-processus ne bloque pas jusqu'à terminaison (ce serait une perte de ressources). Également, la méthode `run()` de `CommunicationThread` fait passer l'état de l'objet à `décrementable` puis termine le fil d'exécution du thread. C'est quand son timer tombe à 0 qu'un nouveau "dumb thread" est créé avec une unique tâche de faire la lecture et le parsing. Cela se justifie par le fait que la méthode qui fait l'appel à la décrémentation (propre à l'objet `Timer` régissant l'application) ne devrait pas effectuer ce travail. En effet, il pourrait perturber la gestion de l'écoulement du temps s'il est trop lourd et effectué dans le même fil d'exécution.

5 Implémentation de l'application

5.1 Justification des choix généraux

5.2 Discussions des choix d'implémentation

5.2.1 Gestion du temps

5.2.2 Exécution des logiciels

5.2.3 Parsing des sorties

5.2.4 Stockage et agrégation des informations

5.2.5 Interactivité avec l'application

5.3 Sélection des logiciels à intégrer

6 Manuel d'utilisation

7 Tests et évaluation de l'application

7.1 Simulation d'équipements

7.2 Déploiement dans un milieu domestique

7.3 Stress-test dans un environnement peuplé

Références

- [1] 6lowpan. Wikipedia article, URL : <https://fr.wikipedia.org/wiki/6LoWPAN>.
- [2] Application firewall. Wikipedia article, URL : https://en.wikipedia.org/wiki/Application_firewall.
- [3] Cable broadband technology gigabit evolution. Dernière lecture le 27/11/2018.
- [4] Cisco asa. Wikipedia article, URL : <https://en.wikipedia.org/wiki/Cisco ASA>, dernière lecture le 19/12/2018.
- [5] Firewall(computing). Wikipedia article, URL : [https://en.wikipedia.org/wiki/Firewall_\(computing\)](https://en.wikipedia.org/wiki/Firewall_(computing)).
- [6] Internet growth statistics par internet world stats. <http://www.ti.com/lit/ml/swrb028/swrb028.pdf>. Dernière lecture le 26/11/2018.
- [7] Les firewalls. Community article, URL : <https://www.frameip.com/firewall/>.
- [8] Pare-feu(informatique). Wikipedia article, URL : [https://fr.wikipedia.org/wiki/Pare-feu_\(informatique\)](https://fr.wikipedia.org/wiki/Pare-feu_(informatique)).
- [9] Security in the Internet of Things. https://www.windriver.com/whitepapers/security-in-the-internet-of-things/wr_security-in-the-internet-of-things.pdf, 2015.
- [10] Alhafidh Basman M.Hasan and William Allen. Design and simulation of a smart home managed by an intelligent self-adaptive system. *Int. Journal of Engineering Research and Application*, 2016.
- [11] Isha Batra and Kr. Luhach Ashish. Analysis of lightweight cryptographic solutions for Internet of Things. *Indian Journal of Science and Technology*, 9, July 2016.
- [12] Dan-Radu Berte. Defining the IoT. *De Gruyter*, 2018.
- [13] Cisco. *CLI Book 2 : Cisco ASA Series Firewall CLI Configuration Guide*, 9.6, May 2018. URLs : <https://www.cisco.com/c/en/us/td/docs/security/asa/asa96/configuration/firewall/asa-96-firewall-config/access-idfw.html>, <https://www.cisco.com/c/en/us/td/docs/security/asa/asa96/configuration/firewall/asa-96-firewall-config/access-idfw.pdf>.
- [14] Jorge Granjal, Edmundo Monteiro, and Jorge Sá Silva. Security for the internet of things : A survey of existing protocols and open research issues. *IEEE Communications Surveys & Tutorials*, pages 1–1, 07 2015.
- [15] Robert J. Shimonski, Debra Littlejohn Shinder, and Thomas W. Shinder, editors. *The Best Damn Firewall Book Period*. Syngress, Burlington, 2003.
- [16] Kalle Kokko. Next-generation firewall case study. Master's thesis, South-Eastern Finland University of Applied Sciences, 2017.
- [17] Sylvain Kubler, Kary Främling, and Andrea Buda. A standardized approach to deal with firewall and mobility policies in the IoT. *Pervasive Mobile Computing*, 2014.
- [18] Vishwas Lakkundi and Keval Singh. Lightweight dtls implementation in coap-based internet of things. 09 2014.
- [19] Engin Leloglu. A review of security concerns in Internet of Things. *Journal of Computer and Communications*, 2017.
- [20] Manuel Leos Rivas. Securing the home IoT network. SANS Institute Reading Room, URL : <https://www.sans.org/reading-room/whitepapers/hsoffice/paper/37717>, 2014.
- [21] Huichen Lin and Neil Bergmann. IoT privacy and security challenges for Smart Home environments. *Information (online MDPI journal)*, 2016.
- [22] Imran Makhdoom, Mehran Abolhasan, Justin Lipman, Ren Liu, and Wei Ni. Anatomy of threats to the internet of things. *IEEE Communications Surveys and Tutorials*, PP, 10 2018.
- [23] Norbert Nthala and Ivan Flechais. Informal support networks : an investigation into Home Data Security Practice. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS)*. USENIX, August 2018.
- [24] Dariusz Palka and Marek Zachara. Learning web application firewall - benefits and caveats. pages 295–308, 01 2011.
- [25] Alaauddin Shieha. *Application Layer Firewall Using OpenFlow*. PhD thesis, University of Colorado, 2014.