UNIVERSITE DE SAAD DAHLAB DE BLIDA -1-

Faculté des sciences

Département d'informatique

MEMOIRE DE MAGISTER

Spécialité : Informatique Répartie et Mobile (IRM)

PROTECTION ET SÉCURISATION DES DONNÉES MÉDICALES SUR LE CLOUD COMPUTING

Par

Lynda KACHA

Devant le jury composé de :

N. BOUSTIA	Maître de conférences (A), U. Blida-1-	Présidente
D. BENNOUAR	Maître de conférences (A), U.Bouira	Examinateur
S. OUKID	Maître de conférences (A), U. Blida-1-	Promotrice
N. BENBLIDIA	Professeur U. Blida -1-	Co-Promotrice

Blida, Novembre 2015

RESUME

Le Cloud Computing est un terme générique désignant toute prestation de services hébergés sur internet, où les utilisateurs peuvent stocker et accéder à distance à leurs données à tout moment et de n'importe où.

Le Cloud Computing est désormais adopté dans divers domaines, en raison de son faible coût, sa haute disponibilité et son évolutivité. La santé est l'un des domaines qui peut bénéficier du Cloud Computing. Le Cloud Computing facilite l'échange d'information entre les professionnels de santé ; ce qui peut améliorer la qualité des soins médicaux et réduire les coûts.

Cependant, pour bénéficier du Cloud Computing, les institutions médicales doivent déplacer les données sensibles de leurs patients vers un serveur tiers. Le déplacement des données vers le Cloud Computing, implique le déplacement du contrôle de ces données vers le fournisseur de service définitivement. Par conséquent, la sécurité et la confidentialité des informations deviennent une question importante.

Ce travail aborde ce problème et propose une solution pour sécuriser les données médicales sur le Cloud Computing. Notre contribution consiste en une architecture de sécurisation qui s'exécute sur une infrastructure Cloud publique ainsi que sur l'infrasturcure privée de l'hôpital.

الملخص

الحوسبة السحابية هي مصطلح عام يشير الى المصادر والخدمات المتوفرة على شبكة الأنترنت والتي تسمـــح للمستعملين التخزين والحصول على المعلومات من أي مكان وفي أي وقت.

الحوسبة السحابية هي متخذة حاليا في مختلف المجالات وذلك بسبب انخفاض تكلفته، توافره العـالي وإمكانيـة تطويره. الصحة هي أحد المجالات التي يمكن ان تستفيد من الحوسبة السحابية. الحوسبـة السحابية تسهل تبـادل المعلومات بين العاملين في مجال الصحة هذا ما يحسن نوعية العناية الصحية مع تخفيض التكاليف.

لكن من اجل الاستفادة من إيجابيات الحوسبة السحابية يجب على المؤسسات الصحية نقل المعلومات الحساسة لمرضاها الى خادم خارجي. هذا ينطوي على نقل السيطرة على هذه المعلومات الى مـزود الخدمة بشكل دائـم وبالتالي يصبح سر وأمن المعلومات مسالة هامة.

مساهمتنـــا تكمن في هندســة أمنيــة والتـي تنفذ في بنيـــة سحابيــة حسابيـــة عموميــة وعلـى البنيـــة الخاصـــة للمؤسسـة.

ABSTRACT

Cloud computing is a generic term referring to any provision of services hosted on the Internet, where users can store and remotely access their data anytime and anywhere. Cloud computing is now adopted in various areas, due to its low cost, high availability and scalability. Healthcare is one area that can benefit from cloud computing. Cloud computing facilitates information's sharing between health professionals which can improve the quality of healthcare and reduce costs.

However, to benefit of cloud computing, medical institutions must move sensitive data of their patients to a third serverur. Moving data to the cloud, involves moving the control of these data to the service provider definitively. Therefore, the security and confidentiality of information becomes an important issue.

This work addresses this issue and proposes a solution to secure medical data on the Cloud Computing. Our contribution consists of a security architecture running on a public Cloud infrastructure as well as on private hospital infrasturcure.

REMERCIEMENTS

J'adresse mes remerciements à toutes les personnes qui m'ont aidée dans la réalisation de ce mémoire.

Je remercie en premier lieu mes encadreurs : Docteur S. Oukid et Professeur N.Benblidia qui ont dirigé mon travail. Je les remercie pour leur aide et leurs conseils.

Je remercie également M^{r} H.Bekkouch pour sa précieuse aide.

Je tiens particulièrement à remercier mon père qui m'a toujours encouragé et soutenu. Sans lui je n'y serais jamais arrivé.

Un grand merci à mon mari pour son soutien. Merci d'être à mes côtés. Avec toute ma tendresse.

Je dédie ce travail à ma mère qui m'a toujours poussé à aller loin et à atteindre mes objectifs. A mon petit frère Mouloud, toute ma famille et ma belle-famille.

LISTE DES ILLUSTRATIONS, GRAPHIQUES ET TABLEAUX

TABLE DES MATIERES

INTRODUCTION

Le coût élevé engendré par la construction et la maintenance de serveurs spécialisés, et les besoins de collaboration entre les institutions médicales ont suscité l'intérêt du secteur médical au Cloud Computing. Le Cloud Computing introduit une nouvelle façon de fournir des services aux communautés médicales. Il permet un accès rapide et ubiquitaire aux informations en offrant une infrastructure d'échange permettant l'intégration et le partage d'informations entre les différentes institutions médicales.

Malgré les avantages que peut apporter le Cloud pour les institutions médicales, les problèmes de sécurité et de confidentialité représentent l'obstacle majeur à l'adoption du Cloud Computing par le domaine médical.

En comparaison avec le modèle informatique traditionnel où les utilisateurs ont un contrôle direct sur les aspects majeurs de la sécurité (car les données résident sur leurs ordinateurs), les utilisateurs Cloud doivent compter sur les fournisseurs de services pour sécuriser leurs données . Cette seule caractéristique soulève cependant, de nombreuses questions en matière de sécurité et de confidentialité.

Bien que les fournisseurs de services Cloud glorifient la sécurité et la fiabilité de leurs services, les déploiements actuels des services Cloud ne sont pas assez sûrs, ni assez fiables qu'ils le prétendent [23,26]. En 2009 par exemple, les principaux fournisseurs Cloud ont été victimes de plusieurs incidents successifs

(Amazon’s Simple Storage Service (S3) a été interrompu 2 fois en Février et en Juillet 2009. Cet incident a entrainé un arrêt dans certains sites du réseau, en Mars (mai) 2009, des failles de sécurité dans Google Docs ont conduit à une fuite d’information privées des utilisateurs. Google Gmail a eu un échec global de plus de 4 h, Microsoft Azure a également subit un grave incident de panne pendant 22h).

Etant donné que le Cloud est un domaine en émergence, il y a un manque de consensus en matière de sécurité et de confidentialité des données [20,25] . Les lois et réglementations sur la protection de la vie privée sont obsolètes dans ce nouvel environnement. Ils ne sont plus applicables à la nouvelle relation entre les utilisateurs et les fournisseurs, relation qui contient désormais trois parties (l’utilisateur du service Cloud, le fournisseur du service Cloud et le fournisseur de l’infrastructure Cloud) [40,26,27]. De plus, l’évaluation et la comparaison entre les différents services Cloud posent un problème pour les utilisateurs novices. [20]

L’objectif général de notre travail consiste à proposer une solution pour sécuriser les données médicales sur le Cloud. Nous avons pour cela adopté le plan suivant :

1- Nous présenterons dans le premier chapitre un état de l’art sur les architectures Cloud. Nous commencerons par présenter le Cloud Computing, nous décrirons son architecture en couches, ses modèles de déploiement, ses modèles de services ainsi que ses principales caractéristiques par rapport à une architecture traditionnelle. Nous finirons ce chapitre par des exemples de quelques produits commerciaux.

2- Le deuxième chapitre sera consacré à la sécurité. Il est composé principalement de deux parties : dans la première partie, nous analyserons les questions de sécurité dans le Cloud Computing d’une manière générale. La

deuxième partie sera consacrée à la sécurité des données dans le Cloud Computing. Nous analyserons, dans cette partie, la sécurité et la confidentialité des données associées au Cloud Computing à travers le cycle de vie des données. Nous metterons l'accent sur les nouveaux problèmes relatifs aux données introduits par l'utilisation du Cloud Computing.

3- Dans le troisième chapitre, nous présenterons notre proposition pour sécuriser les données médicales sur le Cloud. Nous décrirons, en premier lieu, quelques travaux relatifs à notre sujet. Nous analyserons et comparerons ensuite parmis ces travaux, ceux relatifs à notre problématique. Nous présenterons, en dernier lieu, notre architecture et nos motivations.

4- Nous conclurons ce mémoire par le bilan du travail effectué ainsi que quelques perspectives de recherche.

CHAPITRE 1
ETAT DE L'ART SUR LES ARCHITECTURES DE CLOUD COMPUTING

1.1 Introduction

Le Cloud Computing est un terme générique désignant toute prestation de service hébergée sur internet, où les utilisateurs peuvent stocker et accéder à distance à leurs données à tout moment et de n'importe où.

Contrairement au modèle informatique traditionnel où les données et la puissance de calcul des utilisateurs se trouvent sur leurs systèmes informatiques, les utilisateurs ne détiennent plus l'infrastructure physique qui héberge les services, mais louent plutôt l'usage à partir d'un prestataire de service professionnel. Ils consomment les ressources en tant que service et payent uniquement pour les ressources utilisées. De plus, le partage des ressources informatiques entre plusieurs consommateurs réduit significativement les coûts. Les applications Cloud éliminent le besoin d'installer et d'exécuter les applications sur les ordinateurs des utilisateurs, ce qui allège le fardeau de la maintenance logicielle, le suivi des applications et l'entretien. [2,3,10,21]

En outre le Cloud est facile à utiliser car il nécessite peu d'expertise. L'analogie est souvent faite avec l'électricité où les utilisateurs finaux peuvent utiliser les services des fournisseurs sans se préoccuper de la complexité technique de ces systèmes.

En effet, étant donné qu'on est dans une logique de consommation de service, la responsabilité de l'exploitation, du déploiement et de la maintenance de l'infrastructure est à la charge du fournisseur. [3,18]

Dans le Cloud Computing il existe traditionnellement deux types de fournisseurs : [10,26,27]

- le fournisseur de l'infrastructure qui gère la plateforme Cloud et loue les ressources selon un modèle de tarification basé sur l'utilisation.

- les fournisseurs de services qui louent les ressources d'un ou de plusieurs fournisseurs d'infrastructure pout servir les utilisateurs finaux. (cf. Figure 1.1)

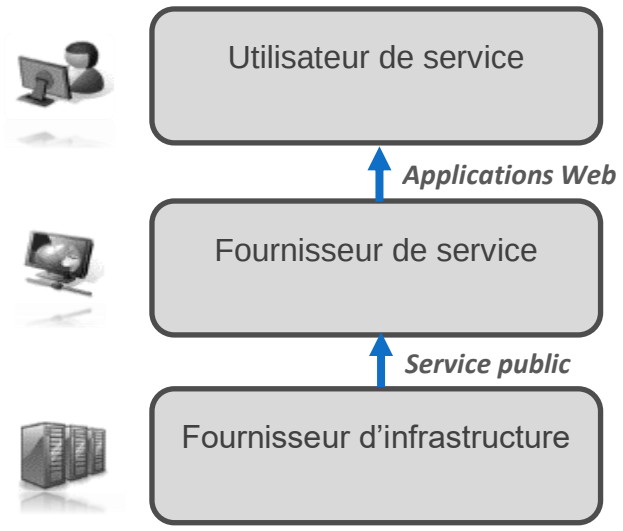

Figure 1.1 : Acteurs du Cloud Computing. [26]

L'émergence du Cloud Computing a eu un impact considérable sur l'industrie informatique au cours de ces dernières années, où de grandes sociétés telles que Google, Amazon et Microsoft s'efforcent de fournir des plateformes Cloud plus puissantes, plus fiables et plus rentables.

1.2 Qu'est-ce que le Cloud Computing ?

Il existe plusieurs définitions du Cloud Computing qui semblent se concentrer uniquement sur certains aspects de cette technologie.

La raison principale de l'existence de plusieurs perceptions du Cloud est que celui-ci, et contrairement à d'autres termes techniques, n'est pas une nouvelle technologie mais plutôt un nouveau modèle de fonctionnement qui regroupe un ensemble de technologies et de concepts existants [10,28,19,21,20]. En effet, la plupart des technologies utilisées dans le Cloud comme la virtualisation et l'informatique utilitaire ne sont pas nouvelles. Le Cloud met à profit ces technologies existantes pour répondre aux exigences technologiques et économiques de la demande actuelle. [10]

La définition la plus utilisée dans la littérature est celle de NIST (National Institut of Standard and Technology) car elle couvre la plupart des aspects essentiels du Cloud. NIST définie le Cloud Computing comme l'ensemble des disciplines technologies et modèles commerciaux, utilisés pour délivrer des capacités informatiques (logiciel, plateformes, matériel) comme un service à la demande.

Le Cloud Computing est donc un modèle qui consiste à proposer les ressources informatiques sous forme de service à la demande et accessibles via internet. Ces services sont généralement divisés en trois catégories : l'infrastructure en tant que service (IaaS), la plateforme en tant que service (PaaS), l'application en tant que service (SaaS).

1.3 Historique

Il n'y a pas de date clé à laquelle on pourrait dire que le Cloud Computing est né. Cependant, l'idée principale derrière le Cloud Computing n'est pas nouvelle. John McCarthy (un chercheur américain en informatique) suggéra, en 1961,

l'utilisation de ressources informatiques comme un utilitaire, permettant de disposer de ressources informatiques comme un service public. [4,10]

Le terme « Cloud » est issu du monde des télécommunications quand les fournisseurs de télécommunication ont commencé à proposer les services de réseau VPN pour le transport des données. Les principes du Cloud Computing sont très similaires, car les fournisseurs proposent un environnement virtuel qui est alloué dynamiquement pour satisfaire les besoins de l'utilisateur. [9,2]

Les années 80 furent aussi le début des concepts de la virtualisation. La virtualisation est la base du Cloud Computing. Cette notion permet une gestion optimisée des ressources matérielles dans le but de pouvoir y exécuter plusieurs systèmes virtuels sur une seule ressource physique.

Amazon a joué un rôle clé dans le Cloud Computing en lançant AWS (Amazon Web Service) en 2006 [11]. Amazon a proposé de louer ses ressource aux entreprises, afin de rentabiliser l'énorme infrastructure prévue pour absorber les charges en périodes des fêtes mais plutôt inutilisée le reste de l'année. Depuis, Amazon investit massivement dans ce domaine et continue d'agrandir son parc et ses services.

Récemment, d'autres acteurs du monde IT comme Google et Microsoft proposent à leur tour des services similaires.

Tous ces concepts ont amené, petit à petit, à inventer une nouvelle manière de proposer l'informatique « comme un service ».

1.4 Technologies annexes

Le Cloud Computing est souvent comparé aux technologies ci-dessous, chacune partageant certains aspects avec celui-ci :

- Grid Computing (Grilles de calcul)

Le Grid Computing est un paradigme de calcul distribué qui coordonne les ressources réseau pour atteindre un objectif de calcul commun ; Il permet le partage, la sélection et l'agrégation d'une grande variété de ressources informatiques délocalisées (comme supercalculateurs, clusters de calcul, systèmes de stockage) et les présente comme une ressource unique et unifiée pour résoudre des calculs sur de grandes quantités de donnée. [10,17]

Le Cloud Computing est similaire au Grid Computing, parce qu'il utilise des ressources distribuées pour atteindre ses objectifs. Cependant le Cloud Computing va plus loin, en s'appuyant sur les techniques de virtualisation à plusieurs niveaux (matériel, plateforme et application), afin de permettre le partage des ressources et le provisionnement dynamique de celles-ci. [10]

- Utility Computing (Informatique utilitaire)

L'informatique utilitaire représente le modèle de provisionnement des ressources à la demande et de facturation basée sur l'utilisation plutôt que sur un taux forfaitaire. Elle permet de louer des ressources informatiques en fonction du besoin [10,26]. Le Cloud Computing peut être perçu comme tel, il adopte un système de tarification basé sur la consommation ce qui permet de maximiser l'utilisation des ressources et minimiser les coûts d'exploitation. [10]

- Virtualisation

« Virtualization is the creation of a virtual version of something, such as operating system, a server, a storage device or network resources » [13]

La virtualisation est une technologie qui fait abstraction des détails matériels. Elle permet de faire fonctionner sur une seule machine (machine hôte) plusieurs systèmes d'exploitation (OS invités) et/ou plusieurs applications séparément les

uns des autres, comme s'ils fonctionnaient sur des machines physiques distinctes.

Une machine virtualisée est communément appelé Machine Virtuelle (VM : Virtual Machine). Les VM sont contrôlées par un logiciel appelé hyperviseur, ou contrôleur de machine virtuelle (Virtual Machine Monitor, VMM). L'hyperviseur masque les véritables ressources physiques de la machine hôte afin de proposer des ressources différentes et spécifiques en fonction des applications. Il simule autant de machines virtuelles que de systèmes d'exploitation souhaités. Ces derniers peuvent fonctionner en parallèle en partageant les mêmes ressources [10,32,29,13]. La figure 1.2 représente une comparaison entre un environnement non virtualisé et un environnement virtualisé.

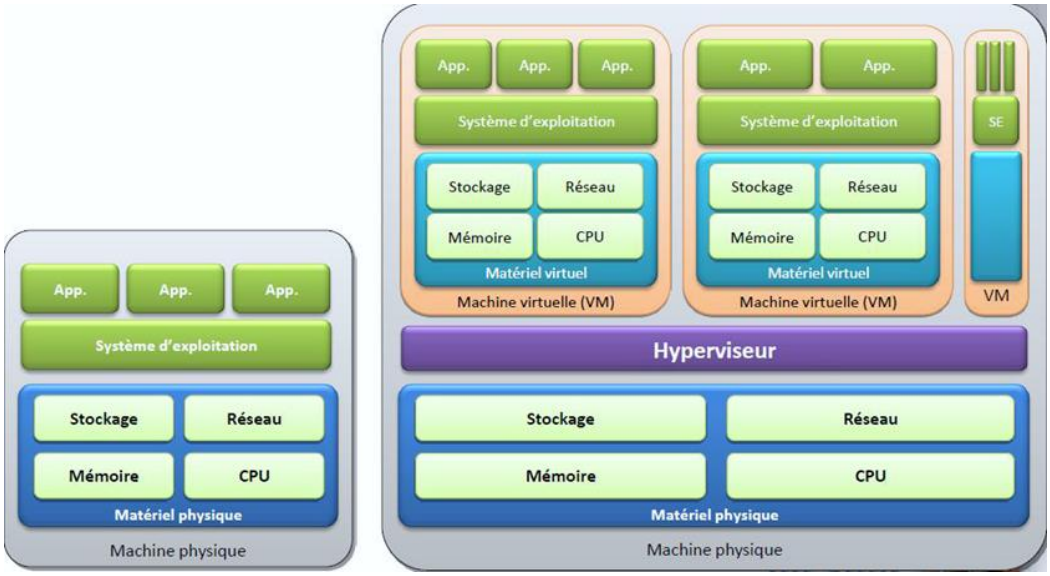

Figure 1.2 : Environnement physique VS Environnement Virtuel. [36]

La virtualisation constitue la base du Cloud Computing, car elle offre la possibilité de mettre en commun des ressources informatiques et l'affectation dynamique des ressources virtuelles aux applications à la demande. [10]

- Automatic Computing (Autogestion)

Elle vise à construire des systèmes informatiques capables de s'auto gérer, c'est-à-dire réagir à des observations internes et externes, sans intervention humaine.

Le but de l'autogestion est de surmonter la complexité de la gestion des systèmes actuels. Bien que le Cloud Computing présente certaines caractéristiques automatiques comme le provisionnement automatique des ressources, son objectif est de minimiser le coût des ressources plutôt que de réduire la complexité du système. [10]

Le Cloud Computing s'appuie donc sur la virtualisation pour fournir des ressources informatiques en tant que service public. Il partage certains aspects avec le Grid Computing et l'automatisation mais diffère par d'autres aspects. [10]

1.5 Architectures de Cloud Computing

Le Cloud Computing transforme la façon dont les organisations perçoivent leurs ressources informatiques. En effet, les organisations évoluent d'une architecture avec un système unique, (composé d'un système d'exploitation unique et d'une application unique), à une architecture Cloud où les ressources sont disponibles à profusion. [5]

Dans le Cloud Computing, les utilisateurs finaux n'ont pas besoin de connaître les spécificités des technologies utilisées pour héberger leurs applications, celles-ci sont entièrement gérées par le fournisseur du service et peuvent être consommées en fonction des besoins des utilisateurs.

Cette section décrit le modèle architectural, commercial et les différents types de Cloud Computing.

1.5.1 L'architecture en couches du Cloud Computing

Les services Cloud sont basés au niveau matériel sur des data Center immenses, et au niveau logiciel sur les techniques de virtualisation offrant ainsi aux utilisateurs des ressources informatiques dont le dimensionnement est variable. [8,10,30].

[10] a présenté l'architecture en couche du Cloud Computing. Celle-ci peut être divisée en 4 couches (cf. Figure 1.3) :

- la couche physique
- la couche infrastructure
- la couche plateforme
- la couche application

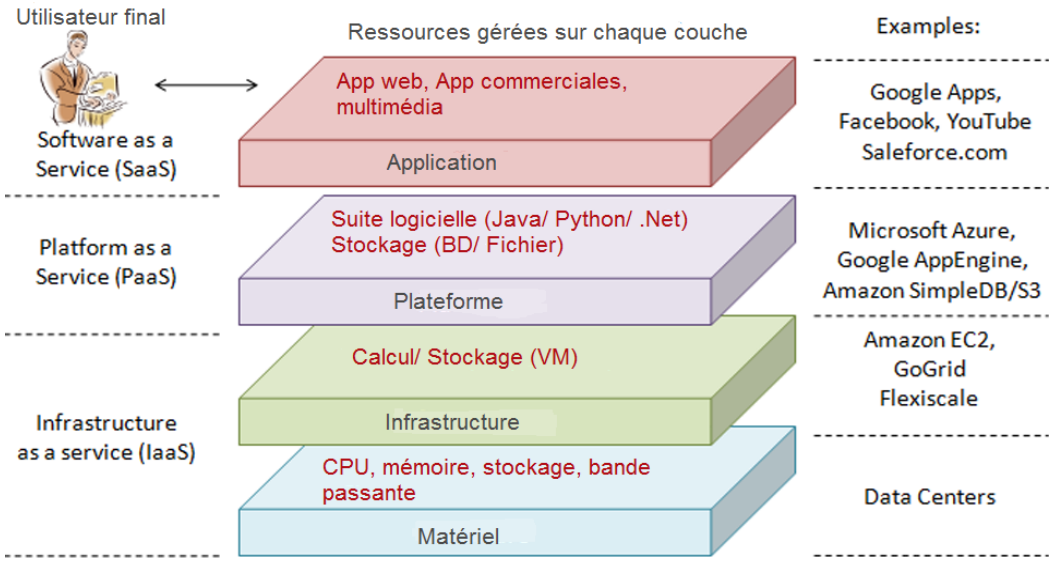

Figure 1.3 : Modèle en couche du Cloud Computing. [10]

- La couche physique

La couche physique est responsable de la gestion des ressources physiques dans le Cloud. Dans la pratique cette couche est généralement implémentée dans des data Center. Un data Center contient généralement des milliers de serveurs organisés en racks et interconnectés par des switches, des routeurs et autres matériels.

Les questions relatives à cette couche incluent la configuration matérielle, la tolérance aux pannes, la gestion du trafic, la gestion des ressources électriques et de refroidissement.

- La couche infrastructure

Aussi connue sous le nom de la couche de virtualisation, cette couche crée un pool de ressources de stockage et de calculs, en partitionnant les ressources physiques, en utilisant les technologies de virtualisation telles que Xen et VMware. La couche infrastructure est un composant essentiel du Cloud, étant donné que des fonctionnalités comme l'affectation dynamique des ressources sont possibles grâce à la virtualisation.

- La couche plateforme

Construite au-dessus de la couche Infrastructure, cette couche consiste en un ensemble de systèmes d'exploitation et d'environnements logiciels. L'objectif de cette couche est d'alléger le fardeau du déploiement des applications directement sur les machines virtuelles.

- La couche application

Au niveau le plus haut de la hiérarchie, cette couche se compose des applications Cloud actuelles. Différentes des applications traditionnelles, les

applications Cloud peuvent tirer parti d'une mise à l'échelle automatique pour atteindre une meilleure performance, disponibilité, et coût d'exploitation.

En comparaison avec des architectures traditionnelles, l'architecture Cloud est plus modulable. Chaque couche est faiblement couplée avec les couches sous-jacentes, ce qui permet à chaque couche d'évoluer séparément, ceci est similaire au modèle OSI pour les protocoles réseaux. L'architecture modulaire permet au Cloud de répondre aux larges exigences des applications tout en réduisant la gestion et l'entretien des frais généraux. [10]

1.5.2 Le modèle commercial du Cloud Computing

Le Cloud Computing utilise un modèle commercial basé sur les services. Toutes les ressources informatiques sont fournies en tant que services à la demande (XaaS).

Conceptuellement chaque couche du modèle décrit précédemment peut être implémentée comme un service pour la couche supérieure [10]. Dans la pratique, le Cloud offre des services qui peuvent être regroupé en trois catégories :

- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)

Ces trois classifications fondamentales sont souvent appelées le modèle SPI, pour Software, Platform et Infrastructure respectivement (cf. Figure 1.4). [22,23]

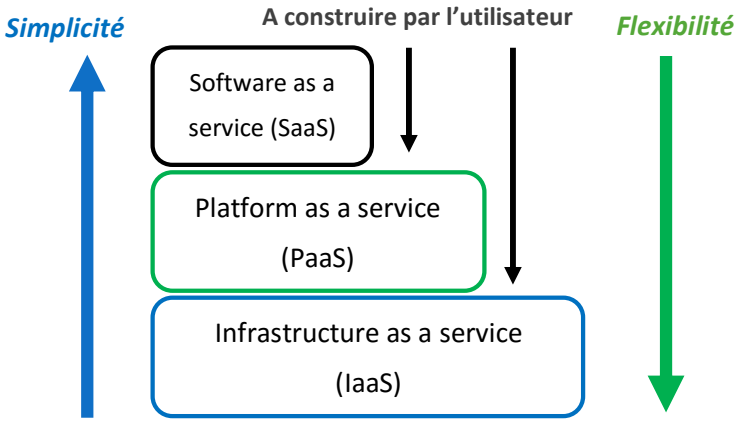

Figure 1.4 : Le modèle SPI du Cloud Computing. [7,24]

- Infrastructure as a Service (IaaS)

Permet de disposer d'une infrastructure à la demande (serveurs, réseaux et stockage) généralement en terme de machines virtuelles. [10,2,9]

Le client n'a pas besoin d'acheter les serveurs requis, le data Center ou les ressources réseau. Il ne paye que pour le temps pendant lequel il utilise le service.

Les clients ne gèrent pas l'infrastructure sous-jacente. Le principal avantage de cette solution réside dans sa flexibilité. [1,12,2,11,9].

- Platfrome as a Service

Fournit un environnement de développement complet, constitué d'un ensemble de logiciels et d'outils de développement hébergés sur l'infrastructure du fournisseur. [10,2]

La PaaS facilite le déploiement des applications sans le coût et la complexité de l'achat et la gestion des couches matérielles et logicielles sous-jacentes. [2,9,12]

- Software as a service

SaaS consiste en un ensemble de logiciels d'exécutant sur la plateforme du fournisseur et délivré à divers clients via internet. [6,10,2,11]

Les utilisateurs n'achètent plus les logiciels, mais les utilisent à la demande, ce qui évite leurs installations et leur mise à jour. [2,9]

Il est possible qu'un fournisseur PaaS exécute ses applications sur l'infrastructure d'un fournisseur IaaS. Cependant, dans la pratique, les fournisseurs IaaS et PaaS font souvent partie de la même organisation (par exemple Google ou Amazon) et sont appelé fournisseurs d'infrastructure ou fournisseurs Cloud. [1]

1.5.3 Les types de Cloud Computing

La migration d'une entreprise vers le Cloud Computing dépend de ses besoins. Quelques fournisseurs de services par exemple mettent en avant la réduction des coûts d'exploitation tandis que d'autres choisissent la sécurité et la fiabilité. Par conséquent, il existe différents types de Cloud chacun avec ses avantages et ses inconvénients.

- Les Clouds publiques

L'infrastructure physique est détenue par le prestataire de service. Les ressources tels que le stockage et les applications sont mises à la disposition des consommateurs via internet, et sont donc hébergées chez le fournisseur et gérer par celui-ci. [2,10,12,11]

Le Cloud public offre plusieurs avantages : entre autre pas de dépenses initiales sur l'infrastructure et le déplacement des risques aux fournisseurs de l'infrastructure. [10,9]

Cependant ce type de Cloud offre moins de contrôle sur les données et les applications. Il est moins sécurisé en comparaison avec d'autres modèles de Cloud Computing étant donné que toutes les applications et les données sont proposées au grand public et accessibles via internet. [1,10]

- Les Clouds privés

Appelé aussi Cloud interne. Dans ce type de Cloud l'infrastructure physique est exploitée exclusivement par une seule organisation.
Ce type de Cloud peut être conçu et géré par l'organisation ou par un prestataire externe. [2,12,11,9,6,10]

Les Cloud privés offrent un degré de contrôle plus élevé sur la performance et la fiabilité. Ainsi la sécurité est renforcée étant donné que seuls les utilisateurs de l'organisation ont accès au Cloud privé.

Les Cloud privés peuvent être comparés à l'intranet et sont souvent critiqués pour être similaire aux serveurs propriétaires traditionnels. [10]

- Les Clouds communautaires

L'infrastructure physique est contrôlée et partagée par plusieurs organisations, et s'appuie sur une communauté d'intérêt. Les membres de la communauté partagent l'accès aux données et aux applications dans le Cloud. L'infrastructure Cloud peut être hébergée chez le fournisseur ou à l'intérieur de l'une des organisations dans la communauté. [11,2,9]

- Les Clouds hybrides

Les Cloud hybrides regroupent deux ou plusieurs infrastructures Cloud distinctes (privée, publiques ou communautaire), et sont liées par une technologie standard permettant la portabilité des données et des applications. [7,2,11]

Ces types de Cloud minimisent les coûts associés aux Cloud privé, et les risques associés aux Cloud publics. En effet, généralement dans ce modèle, le Cloud privé est connecté à un ou plusieurs services Cloud externes. C'est une manière plus sécurisée pour contrôler les données et les applications, et permettre l'accès à l'information via internet. Il permet à l'organisation de servir ses besoins dans le Cloud privé et de se déplacer dans le Cloud public quand certaines nécessités occasionnelles surviennent. [9]

Les applications avec moins d'exigences de sécurité peuvent ainsi être externalisées vers le Cloud public, tout en gardant les services critiques et les données confidentielles dans un Cloud sécurisé et privé sous contrôle. [2]

Pour la plupart des fournisseurs de services, la sélection du modèle approprié dépend du scénario commercial ; Par exemple les applications de calculs scientifiques intensifs sont mieux déployées sur les Cloud publics pour leur rentabilité.

1.6 Caractéristiques du Cloud Compting

Le Cloud Computing offre plusieurs fonctionnalités différentes de l'informatique traditionnelle :

- Multi-location

  Dans un environnement Cloud, les services sont détenus par plusieurs fournisseurs qui partagent la même infrastructure physique.

  Les problèmes liés à la performance et à la gestion des services sont partagés entre les fournisseurs de service et les fournisseurs d'infrastructure. L'architecture en couche du Cloud Computing fournit une division naturelle des responsabilités. Le propriétaire de chaque couche se concentre sur les objectifs associés à celle-ci. [10]

  Cependant, la multi-location introduit également des problèmes concernant de sécurité et de confidentialité des données. [4] [10]

- Partage des ressources

  Les ressources du fournisseur sont mises en commun pour servir plusieurs consommateurs, en utilisant un modèle multi-tenant avec différentes ressources physiques et virtuelles affectées dynamiquement en fonction de la demande des utilisateurs. [11,31,2,25]

  Le partage des ressources permet la réduction des coûts. Il facilite également la maintenance des applications Cloud étant donné qu'elles n'ont pas besoin d'être installées sur chaque machine. [9]

- Géo-distribution et accès ubiquitaire au réseau

  Dans le Cloud Computing, les utilisateurs accèdent aux données et aux applications à l'aide d'un navigateur web, indépendamment de l'appareil utilisé. [9,10,5]

- Orionté service

  Le Cloud adopte un modèle d'exploitation axé sur les services. Chaque fournisseur offre ses services en accord avec le niveau de SLA négocié avec ses clients. Le SLA (Service Level Agreement) est un contrat légal entre un fournisseur et son client. Celui-ci contient par exemple le service délivré et les responsabilités de sécurisation. [1, 10]

- Provisionnement dynamique des ressources

  L'une des principales caractéristiques du Cloud est que les ressources informatiques peuvent être adaptées à la demande de l'utilisateur. En comparaison avec le modèle traditionnel qui fournit les ressources en fonctions des pics de la demande, le provisionnement dynamique des ressources permet l'acquisition des ressources en fonction de la demande actuelle. Ce qui permet de minimiser les coûts, étant donné que les ressources sont fournis par un tiers, et n'ont pas besoin d'être détenue que pour les tâches occasionnels. [2,10,31]

  Les capacités disponibles paraissent souvent illimitées pour les consommateurs et peuvent être affectées dans n'importe quelle quantité et n'importe quand. [11,25]

- Autogestion

  La gestion automatisée des ressources fournit une agilité élevée qui permet aux fournisseurs de services de réagir rapidement aux changements de la demande. Le consommateur peut se provisionner en ressources au besoin et automatiquement sans l'intervention du fournisseur. [10,5,11]

- Payement à l'usage

  Le Cloud Computing utilise un modèle de tarification fondé sur l'usage c.-à-d. que les capacités sont payées à la consommation. Le schéma exact de tarification peut varier d'un service à un autre. [2,10,25]

  La tarification à l'usage diminue le coût d'exploitation. Cependant elle augmente la complexité de la gestion des coûts d'exploitation. [10]

- Sécurité

  Celle-ci qui peut être meilleure que dans les systèmes traditionnels. Les fournisseurs ont souvent plus d'expertise, et sont mieux équipés pour gérer les problèmes de sécurité. Cependant la sécurité reste une préoccupation importante quand les données sont confidentielles. [9]

1.7 Quelques produits commerciaux

1.7.1 Le Cloud Amazon

AWS (Amazon Web Sevice) est un ensemble de service Cloud fournissant le calcul, le stockage et d'autres fonctionnalités qui permettent aux utilisateurs de déployer des applications et des services à la demande.[10] Les offres AWS sont accessibles au public depuis 2006. [33]

- Amazon EC2 (Amazon Elastic Compute Cloud)

  Amazon EC2 est un service Web qui fournit une capacité de calcul redimensionnable dans le Cloud. [14]

EC2 permet aux utilisateurs Cloud de lancer et de gérer des instances de serveur en utilisant des outils disponibles d'Amazon. Les Instances EC2 sont des machines virtuelles fonctionnant au-dessus du moteur de virtualisation Xen . [10]

L'interface EC2 permet d'obtenir et de configurer des capacités avec un minimum d'efforts. Elle fournit un contrôle complet des ressources informatiques et permet d'exécuter les applications sur l'environnement informatique d'Amazon qui a fait ses preuves. [14]

- Amazon S3 (Amazon Simple Storage Service)

Amazon S3 est un service de stockage pour Internet, destiné aux développeurs. Amazon S3 offre une interface de services Web qui permet de stocker et d'extraire des données à tout moment et depuis n'importe où. Il permet aux développeurs d'accéder à une infrastructure hautement évolutive, fiable, sûre, rapide et peu coûteuse. [14]

1.7.2 Le Cloud Microsoft

Microsoft a lancé sa première version de Windows Azure : Windows Azure Beta en 2009, cette dernière a été commercialisée en 2010 et ne cesse de se développer depuis. [33]

- Microsoft Windows Azure

Windows Azure est une plateforme Cloud ouverte et flexible, qui permet de construire, déployer et gérer rapidement des applications à travers un réseau global de data Center gérés par Microsoft. La plateforme Windows Azure permet de créer des applications à l'aide de n'importe quel langage. Les

applications Cloud public peuvent être intégrées avec l'environnement informatique existant. [15]

1.7.3 Le Cloud Google

En 2008, Google a lancé son Cloud Public orienté pour les services web offrant une plateforme (PaaS) nommée « Google App Engine » et permettant l'hébergement d'applications Python et java, ainsi que des applications SaaS regroupées dans la gamme « Google App ». [33]

- Google App Engine

Google App Engine est une PaaS. Elle permet d'exécuter des applications Web sur l'infrastructure Google. Les applications App Engine sont Faciles à développer et à gérer grâce à leur caractère évolutif. Celles-ci s'adaptent aux besoins en termes de trafic et de stockage des données. [16]

- Google App

C'est une suite d'outils de productivité, qui permet à des utilisateurs (une entreprise, un établissement d'enseignement, une collectivité,…) de se connecter et de travailler sur n'importe quel appareil et n'importe où. Google Apps est très facile à utiliser, elle facilite le travail en groupe en partageant les documents sur le Cloud et en les modifiant en temps réel [16].

Les principales applications en ligne sont : Gmail, Google Agenda, Google Drive (pour le stockage et le partage des fichiers), Documents texte, Feuilles de calcul …

1.7.4 Le Cloud SalesForce.com

Créé en 1999, SalesForce.com est un éditeur de logiciel, devenu l'un des pionniers du modèle SaaS, notamment grâce à son outil historique de CRM. SalesForce.com héberge des applications d'entreprise. SalesForce.com ne cesse de se développer depuis. [34]

1.8 Conclusion

Le Cloud Computing offre divers services, permettant à des utilisateurs de gérer leurs activités, stocker leurs données et utiliser des services Cloud sans avoir à investir dans de nouvelles infrastructures, de licences, ou de formation du personnel.

Cependant le stockage de grandes quantités de données, notamment des données sensibles dans le Cloud Computing, pose de nombreux problèmes. Particulièrement, les problèmes liés à la sécurité et à la confidentialité de l'information. [4]

CHAPITRE 2
ETAT DE L'ART SUR LA PROTECTION ET SECURISATION DES DONNEES

2.1 Introduction

Le Cloud Computing s'appuie sur plusieurs technologies existantes, telles que: la virtualisationet les services web qui contribuent à son évolution, mais se confronte également à leurs défis. Le principal défi concerne la sécurité des données et des services fournis.

2.2 Les questions de sécurité dans le Cloud Computing

[77] définit la sécurité du Cloud Computing comme « un sous domaine du Cloud Computing en relation avec la sécurité informatique. Elle implique des concepts tels que la sécurité des réseaux, du matériel et des stratégies de contrôle déployées afin de protéger les données, les applications et l'infrastructure associée au Cloud Computing».

Les études réalisées par l'IDC[1] en 2008 et 2009 (cf. Figure 2.1 et Figure 2.2) sur les risques liés au Cloud Computing , placent la sécurité au premier plan, et c'est le défi numéro un pour les utilisateurs Cloud. [23,21,32,46,26]

[1] International Data Corporation est un organisme spécialisé dans les études de marché, les analyses sur les technologies de l'information, les télécommunications et les technologies pour les consommateurs

En effet, le déplacement des applications et des données sensibles dans un environnement public et partagé est une préoccupation importante pour la plupart des sociétés qui sont familières avec le modèle sur-site traditionnel où les données et les applications résident sur leurs locaux. Par conséquent, il y a un manque d'inconfort quand il y absence de contrôle sur la sécurité. [19,38,25]

Figure 2.1: IDC 2008.

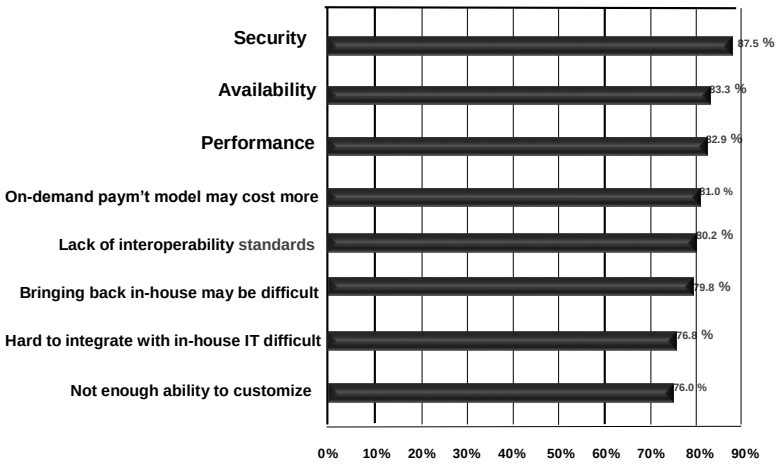

Figure 2.2: IDC 2009.

Pour apaiser ces craintes, les fournisseurs doivent s'assurer que les consommateurs continuent à avoir les mêmes contrôles de sécurité que dans leurs environnements traditionnels. Ils doivent également fournir des preuves aux consommateurs que leurs données et applications sont sécurisées, qu'ils peuvent répondre à leurs SLA et prouver la conformité aux auditeurs. [19,38,25]

D'après [73], il est important de faire la distinction entre les risques de sécurité liés à toutes infrastructures informatiques, et ceux introduits par l'utilisation du Cloud. Ces risques sont généralement associés aux environnements ouverts et partagés. Par conséquent, lors de l'analyse des risques, il est important de séparer les problèmes déjà existants sur les infrastructures médicales de ceux soulevés par le Cloud. Nous traiterons dans ce papier uniquement les problèmes introduits par l'utilisation du Cloud dans l'e-santé.

En raison des modèles de services employés, des modèles opérationnels et des technologies utilisées, le Cloud peut présenter des risques supplémentaires par rapport à une architecture traditionnelle. En effet :

- les applications et les données sur la plateforme Cloud n'ont pas d'infrastructure fixe, ni de périmètre de sécurité. Dans le cas d'une violation de la sécurité, il est difficile d'isoler une ressource physique particulière qui est menacée, ou qui a été compromise. [23]
- les applications et les données sont stockées sur l'infrastructure du fournisseur. Il y a donc une perte de contrôle sur les questions relatives à la sécurité et à la confidentialité des données.
- en raison des problèmes de transparence, aucun fournisseur Cloud ne permet à ses clients de mettre en œuvre des systèmes de contrôles. [21,45,32]
- selon les modèles de services délivrés du Cloud Computing, les services Cloud peuvent être détenus par de multiples fournisseurs. Comme il y a un conflit d'intérêts, il est difficile de déployer des mesures de sécurité unifiées. [23]
- en raison de l'ouverture du Cloud Computing et du partage des ressources virtualisées par les multi-tenants, des utilisateurs non autorisés peuvent accéder

à des données utilisateurs. Ce qui constitue une menace directe sur la confidentialité et la sécurité des données et des services Cloud. [23,26,21]

- étant donné que la plateforme Cloud doit faire face à un stockage massif d'information et fournir un accès rapide, les mesures de sécurité doivent répondre au traitement massif d'information. [23]
- les fournisseurs de services établissent souvent un SLA (Service Level Agreement) pour souligner la sécurité et la confidentialité des services relatifs. Il y a un manque d'une méthodologie standard pour concevoir un SLA. Le SLA fournit des services et des dérogations. Celles-ci n'aident pas vraiment les consommateurs à compenser leurs pertes. [21]

2.2.1 Les Objectifs de la sécurité dans le Cloud Computing

Selon [26] , il existe cinq objectifs pour atteindre une sécurité adéquate : la disponibilité, la confidentialité, l'intégrité des données, le contrôle et les audits. Les 5 objectifs sont intégrés automatiquement et aucun d'eux ne doit manquer pour atteindre une sécurité adéquate. Néanmoins, peu de systèmes Cloud actuels peuvent atteindre les 5 objectifs en même temps.

- La disponibilité

L'objectif de la disponibilité dans les systèmes Cloud est d'assurer que les utilisateurs bénéficient des ressources et des services en permanence. L'architecture Cloud doit garantir un accès aux services à très haute disponibilité et sans interruption [47]. En raison de sa nature web-native, le système Cloud permet à ses clients d'accéder aux données et aux applications à partir de n'importe quel endroit. [26,44]

Dans les environnements informatiques traditionnels, la principale menace pour la disponibilité des données provient des agressions extérieures. Cependant, dans le Cloud, il existe deux principales menaces:

- *les attaques externes* ***:*** une personne non autorisée peut accéder à des données sensibles, car le contrôle n'est pas aux mains des propriétaires. (incluent toutes les menaces dans les scénarios impliquant des infrastructures publiques). [42,27]

- *le fournisseur de service* lui-même constitue une menace, car les données sont maintenues dans ses locaux. [27]

De plus, l'indisponibilité des services Cloud peut être causés par d'autres problèmes tels que : des dysfonctionnements des réseaux, des erreurs matérielles/logicielles et les catastrophes naturelles.

➢ La confidentialité des données/services

La confidentialité consiste à assurer que seules les personnes autorisées puissent accéder aux ressources Cloud.

La confidentialité dans les systèmes Cloud est un grand obstacle à l'adoption de celui-ci. Actuellement, les offres Cloud sont essentiellement publiques et donc exposées à plus d'attaques, comparées à celles hébergées sur les Data Center privés. [26,32,30]

➢ L'intégrité des données/services

L'intégrité des données dans le Cloud signifie protéger l'information des modifications non autorisées. Assurer l'intégrité est une tâche fondamentale. [26,23]
L'intégrité des données est facile à réaliser dans un système autonome avec une seule base de données. Cependant dans le Cloud Computing le problème de l'intégrité des données s'amplifie. En effet le plus grand défi est la gestion des transactions à travers de multiples sources de données. Les accès

concurrents des utilisateurs ne doivent pas compromettre la cohérence des données. [44,47]

De plus, les données sont stockées dans plusieurs endroits. Leur récupération consomme de la bande passante réseau, et certains fournisseurs Cloud comme Amazon obligent leurs utilisateurs à payer les frais de transfert. Comment vérifier directement l'intégrité des données dans le Cloud, sans avoir télécharger les données au préalable ? [23,39]

- Le contrôle

Le contrôle dans le Cloud signifie réglementer l'utilisation du système (l'infrastructure, les applications et les données).

Le Cloud Computing implique des calculs distribués sur de multiples ensembles de données à grande échelle, ainsi que l'accès aux données à travers internet. Par conséquent, effectuer des calculs distribués dans le Cloud sur de telles données, particulièrement quand celles-ci sont confidentielles, soulève de nombreux problèmes de sécurité et de confidentialité [26]. L'utilisateur doit s'assurer que le fournisseur produise des résultats valables quand celui-ci fait des calculs sur les données hébergées chez lui. [32]

- Les audits

Les audits signifient vérifier ce qui se passe dans le Cloud. Les fournisseurs doivent pouvoir être audité sur la sécurité de leurs infrastructures et de leurs solutions par les organisations clientes. Ces audits doivent suivre la classification des exigences auxquelles ils doivent s'y tenir, comme les contrats des clients, les lois et les règlements. [55,22]

2.2.2 Les dimensions de la sécurité dans le Cloud Computing

Les organisations utilisent le Cloud Computing dans une variété de modèles de service (SaaS, PaaS et IaaS) et modèles de déploiement (privé, public, hybride et communautaire). Par conséquent, les risques sont différents selon le niveau du Cloud utilisé, en effet, selon que l'utilisateur dépende d'un logiciel, d'une plateforme, ou d'une infrastructure, la gestion de sécurité sera différente, comme le montre la figure suivante :

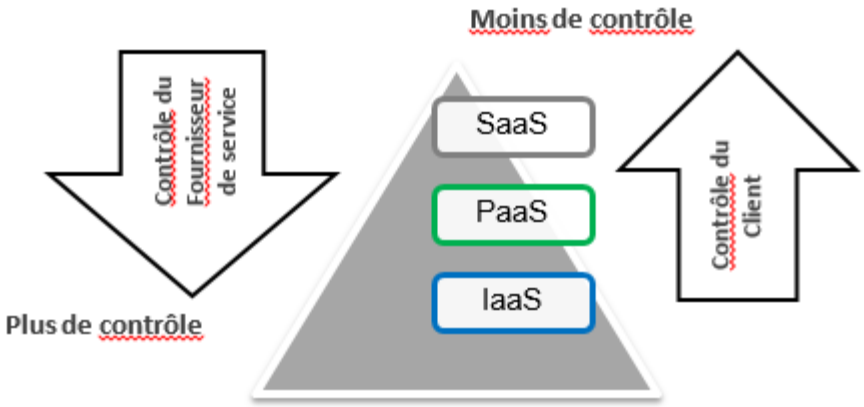

Figure 2.3: Contrôle de sécurité dans le Cloud. [24]

La sécurité de l'infrastructure dépendra également de la manière dont celle-ci est déployée: publique, privée ou communautaire. En effet, si le contrôle de la sécurité sur un Cloud privé est à priori élevé puisque maitrisé, le niveau de contrôle sur un Cloud public est considérablement plus bas.

Bien que les préoccupations liées à la sécurité dans le Cloud Computing peuvent être regroupées en un certain nombre de dimensions (13 selon CSA[2] -

[2] 13 domaines d'intérêts groupés en trois sections : l'architecture Cloud, la gouvernance dans le Cloud, le fonctionnement dans le Cloud.

Cloud Security Alliance- , 7 selon Gartner [3]), ces dimensions ont été regroupées dans ce mémoire en trois grands domaines (les problèmes relatifs à SaaS, les problèmes relatifs à PaaS et les problèmes relatifs à IaaS). Etant donné que le Cloud Computing est une technologie en couche, les mécanismes de sécurité doivent être réalisés sur chaque couche du modèle SPI (Application, Plateforme et Infrastructure).

2.3 La sécurité sur les différentes couches du modèle SPI

Le Cloud Computing utilise trois modèles de prestation, par lesquelles différents types de services sont fournis à l'utilisateur final. Ces modèles de services mettent aussi différents niveaux d'exigences en matière de sécurité dans le Cloud.[44]

Cette partie explore les différentes couches du Cloud Computing, ainsi que leurs problèmes de sécurité associés. Elle s'appuie principalement sur les travaux de [19] et [44] .

2.3.1 problèmes de sécurité relatifs à SaaS

Dans le modèle SaaS, les données de l'entreprise sont stockées sur les Data Center des fournisseurs SaaS, avec les données d'autres entreprises. Le fournisseur SaaS peut héberger les applications sur son propre serveur, ou les déployer sur une infrastructure Cloud fournie par un fournisseur tiers comme Amazon, ou Google. Le fournisseur Cloud pourrait, en plus, répliquer les données sur de multiples endroits à travers divers pays pour un haut degré de disponibilité. [44]

[3] Accès privilégié des utilisateurs, conformité, localisation des données, ségrégation des données, récupération des données, soutien aux enquêtes, viabilité à long terme.

Avec SaaS, le client doit dépendre du fournisseur pour les mesures de sécurité appropriées. L'adoption des applications SaaS peut soulever les problèmes suivants :

- La sécurité des applications

  SaaS est un logiciel déployé sur internet. La principale caractéristique de ce modèle est que sa gestion se fait à partir d'un emplacement central, plutôt que sur chaque site client. [19,44]

  Etant donné que les applications web et SaaS sont fortement couplées dans la fourniture de services aux utilisateurs, la plupart des menaces de sécurité sur les applications web sont également posées par le modèle SaaS. [19, 44]

- La multi-location

  La multi-location est l'une des caractéristiques principales du Cloud Computing, dans laquelle les données de plusieurs locataires sont susceptibles d'être stockées sur le même emplacement physique. Ainsi, le risque de fuite des données entre les locataires est élevé. Les politiques de sécurité doivent assurer que les données des concurrents soient séparer les uns des autres. [19,44,25,21,23,26]

- La sécurité des données

  La sécurité des données est une préoccupation commune à toutes les technologies, mais elle devient un défi majeur, quand les utilisateurs SaaS comptent sur les fournisseurs pour sécuriser leurs données. [19,44,24,32]

- L'accessibilité

L'accessibilité aux applications web et aux données via internet par un navigateur web, rend l'accès à partir de n'importe quel périphérique réseau facile, y compris des ordinateurs publics et appareils mobiles. Cependant elle expose aussi le service à des risques de sécurité additionnels. [19,26]

Les questions relatives à l'accès aux données sont principalement liées aux politiques de sécurité fournis aux utilisateurs quand ceux-ci accèdent aux données. Les politiques de sécurité d'une entreprise peuvent empêcher certains employés d'accéder à un certain type de données. Ces politiques de sécurité doivent être respectés par les fournisseurs pour éviter les intrusions aux données par les utilisateurs non autorisés. Le modèle SaaS doit être suffisamment souple pour intégrer les politiques spécifiques des organisations [44]. Le contrôle d'accès dans le Cloud est donc plus important que jamais, étant donné que le Cloud et toutes ses données sont accessibles via internet. [22]

2.3.2 Les problèmes de sécurité relatifs à PaaS

PaaS facilite le déploiement des applications Cloud, sans l'achat et le maintien des couches matérielles et logicielles sous-jacentes. La sécurité des applications comprend deux couches logicielles [19]: la sécurité de la plateforme PaaS elle-même (c.-à-d. le moteur d'exécution) et la sécurité des applications utilisateurs déployées sur la plateforme.

Du point de vue développement des applications, les développeurs se confrontent à la complexité de concevoir des applications sécurisées qui peuvent être hébergées dans le Cloud. La vitesse à laquelle les applications changent dans le Cloud affecte la sécurité. Les applications PaaS doivent être mises à jour régulièrement. Les développeurs doivent, donc, veiller à ce que les processus de

développement de leurs applications soient suffisamment souples, pour s'adapter aux changements. Cependant, tout changement dans les composants PaaS peut compromettre la sécurité de leurs applications. [19]

Dans le modèle PaaS, les développeurs n'ont généralement pas accès aux couches sous-jacentes, et n'ont aucune assurance que les outils de développement de l'environnement fournit par le fournisseur PaaS soient sécurisé. [19]

2.3.3 Les problèmes de sécurité relatifs à IaaS

IaaS founit un pool de ressources telles que : les serveurs de stockage, les réseaux et les autres ressources informatiques sous la forme de systèmes virtualisés, qui sont accessibles via internet. Avec IaaS, les utilisateurs Cloud ont un meilleur contrôle sur la sécurité par rapport aux autres modèles. Ils contrôlent les applications s'exécutant sur leurs VM et sont responsables de configurer les politiques de sécurité. Cependant l'infrastructure sous-jacente est contrôlée par les fournisseurs Cloud sur qui les fournisseurs IaaS doivent dépendre pour la sécurité. [19,44].

La virtualisation augmente la surface d'attaque, à cause de la couche supplémentaire (couche virtuelle) qui doit être sécurisée (cf. Figure 2.4). Les environnements virtualisés sont vulnérables à tous les types d'attaques pouvant toucher les infrastructures physiques. [19,20,22,32,2,29]

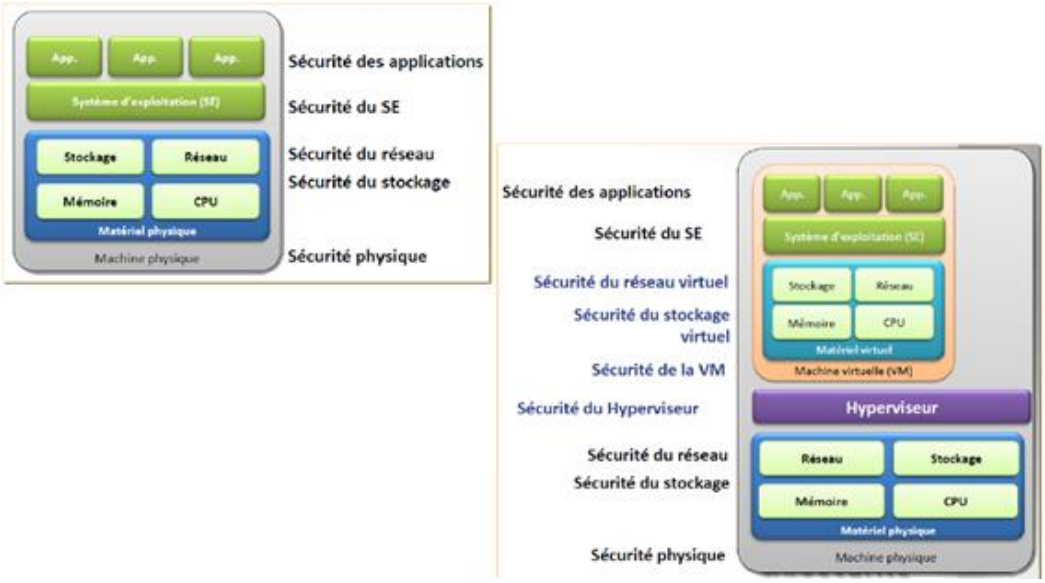

Figure 2.4: Augmentation de la surface d'attaque avec la virtualisation. [36]

Les questions de sécurité de ce modèle comprennnent la sécurité de l'infrastructure physique et la sécurité de l'infrastructure virtuelle:

2.3.3.1 La sécurité de la couche virtuelle

- La sécurité des machines virtuelles

  La sécurité de la VM devient aussi importante que la sécurité de machine physique, et toute faille dans l'une peut affecter l'autre.

  Dans les environnements IaaS, une image VM est un modèle contenant les fichiers de configuration qui sont utilisés pour créer les machines virtuelles. Ainsi ces images sont fondamentales pour la sécurité globale du Cloud. [19]

  Une VM peut être instanciée en utilisant une image malicieuse. Ces images peuvent être le point de départ pour la prolifération des malwares, en injectant du code malveillant dans les autres VM. [19]

De plus, la création, la réplication ou la migration des VM peut exposer leurs contenus au réseau, ce qui peut compromettre la confidentialité et l'intégrité de leurs données.

- L'hyperviseur

L'hyperviseur ou contrôleur de machine virtuelle (VMM) est responsable de l'isolation des VM, par conséquent, si celui-ci est compromis, ses VM peuvent potentiellement être compromises aussi [19,20,2,29]. L'hyperviseur est un logiciel de bas-niveau qui contrôle et surveille les VM ; ainsi comme tout logiciel traditionnel, il comporte des failles de sécurité. En outre, la virtualisation introduit la possibilité de migrer les VM entre les serveurs physiques pour une meilleure tolérance aux pannes, un équilibrage de la charge, ou la maintenance. Cette fonctionnalité utile, peut aussi augmenter les problèmes de sécurité. Un attaquant peut compromettre le module de migration dans l'hyperviseur et transférer une VM victime vers un serveur malveillant. [19,21]

- Les ressources partagées

Les VM situées sur le même serveur peuvent partager le CPU, la mémoire, les E/S… le partage des ressources entre les VM peut introduire des risques. Par exemple : une VM malveillante peut déduire des informations concernant d'autres VM à travers la machine partagée, ou d'autres ressources partagées sans avoir besoin de compromettre l'hyperviseur. En utilisant des canaux secrets, deux VM peuvent communiquer en contournant toutes les règles définies par le module de sécurité de l'hyperviseur. Ainsi, une VM malveillante peut surveiller les ressources partagées sans se faire remarquer par son hyperviseur. [19,26,23,21]

- Les réseaux virtuels

Les composants réseaux sont partagés par différents locataires en raison de la mutualisation des ressources. Les réseaux virtuels augmentent la connectivité

entre les VM, mais introduisent d'importants problèmes en matière de sécurité. [22,19,21]

2.3.3.2 La sécurité de la couche physique

La sécurisation, à ce niveau, consiste à assurer la protection physique des installations informatiques contre les risques d'origine naturelle, ou humaine. Pour cela la conception sécurisée des data center doit être prise en compte (emplacement d'une salle informatique, sécurisation de la partie énergie, climatisation…), ainsi que le contrôle des accès physiques (personnel interne et prestataires externes) [2]. Les zones les plus sensibles doivent être bien délimitées [50]. (cf. Figure 2.5)

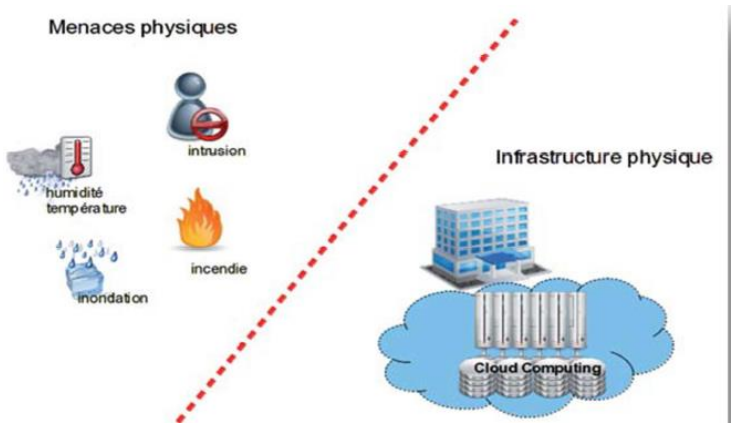

Figure 2.5: Délimitation de l'infrastructure physique. [50]

PaaS ainsi que SaaS sont hébergées au-dessus de IaaS. Ainsi n'importe quelle faille dans IaaS, affectera la sécurité de SaaS et de PaaS. De plus, PaaS offre une plateforme pour créer et déployer des applications SaaS, ce qui augmente la dépendance de sécurité entre eux. Comme conséquence à ces dépendances profondes, toute attaque à n'importe quelle couche des services Cloud peut compromettre les couches supérieures et inversement. Chaque modèle de service comporte ses propres failles de sécurité inhérentes. Cependant ils partagent aussi certains défis qui touchent l'ensemble. [19]

En outre, un fournisseur SaaS peut louer un environnement de développement d'un fournisseur PaaS qui peut à son tour louer une infrastructure d'un fournisseur IaaS. Chaque fournisseur est responsable de sécuriser ses propres services ce qui peut entrainer une combinaison incohérente des modèles de sécurité. Cela créé également une confusion sur les responsabilités de chaque fournisseur une fois qu'une attaque se produit. [19]

2.3.4 Autres risques

- Perte de gouvernance

En utilisant les infrastructures Cloud, le client doit nécessairement céder le contrôle au fournisseur Cloud sur un certain nombre de questions qui peuvent affecter la sécurité [22,40,32,46]. Dans certains cas, il peut être difficile pour le client, dans son rôle en tant propriétaire des données, de vérifier efficacement la gestion de ses données par le fournisseur, et être sûr que ses données sont traitées d'une manière légale [22,19,20].

En outre, le Cloud Computing crée un autre niveau d'utilisateurs privilégiés : les administrateurs Cloud. Ces derniers ont un contrôle total sur les ressources et peuvent donc accéder à toutes les données stockées et les services. Ce type de risque est difficile à détecte car les activités malveillantes des administrateurs sont souvent non détectées, du fait qu'elles sont considérées comme des activités légales par le système. [32,19]

- Le format propriétaire

Il y a actuellement peu d'offres concernant les outils, les processus, les formats standards des données, ou des interfaces des services, susceptibles de garantir la portabilité des données et des applications. Cela peut rendre difficile [22,39,46] :

- la migration d'un client d'un fournisseur à un autre.
- l'utilisation de services de plusieurs fournisseurs.
- la migration des données et des services vers une infrastructure internes.

➢ Fiabilité de la plateforme

La plupart des plateformes Cloud déclarent que la fiabilité atteint 99,9%, mais un taux d'échec de 0,01% peut être désastreux pour l'utilisateur, donc la sauvegarde des données est un aspect important pour faciliter la reprise en cas d'incident. [39,19,44]

➢ La fiabilité du fournisseur

- l*a viabilité à long terme du fournisseur* : L'utilisateur doit s'assurer de la disponibilité de ses données hébergées chez un fournisseur, lorsque celui-ci quitte le marché par exemple. [49,23,24,55]

- l*a nature transitive du fournisseur* : le fournisseur pourrait lui-même sous-traiter ses services par un autre fournisseur sur lesquels l'utilisateur a encore moins de contrôle, et qui doit être également digne de confiance. [32,19]

➢ Les questions juridiques

Les ressources physiques hébergeant les applications et les données dans le Cloud ne sont pas fixes, et évoluent dans le temps. Les données et les applications sont répartis géographiquement sur plusieurs data center dans divers pays ce qui soulève de questions d'ordre juridiques. Les questions juridiques réfèrent aux législation et aux réglementations sur la confidentialité/vie privée dans les pays qui utilisent le Cloud. [71,46,50]

En effet, les fournisseurs ont le droit de divulguer les informations des utilisateurs dans le respect des lois et des demandes des autorités des différents pays où résident les données. [51]

2.4 Analyse de la sécurité dans le Cloud Computing

Dans le Cloud Computing, les vulnérabilités existantes, les menaces et les attaques associées soulèvent plusieurs problèmes de sécurité. Elles affectent directement ou indirectement la confidentialité, l'intégrité et la disponibilité des ressources Cloud, ainsi que les services sur les différentes couches.

[21] et [19] ont analysé les vulnérabilités et les menaces dans le Cloud Computing. Cette partie consiste en un résumé des principales vulnérabilités et menaces présentées de ces travaux. Elle décrit notamment quelques attaques potentielles dans le Cloud Computing .Pour chaque menace et vulnérabilité les modèles de services Cloud affectés par ces problèmes de sécurité sont identifiés par les auteurs. Celles-ci sont résumées dans le tableau 2.1 et le tableau 2.2.

2.4.1 Les vulnérabilités

Une vulnérabilité peut être définie comme une faille dans l'architecture de sécurité du Cloud qui peut être exploité par un adversaire pour accéder aux ressources de l'infrastructure. [19,21]

Le tableau 2.1 représente les principales vulnérabilités dans le Cloud Computing. Il contient une brève description des vulnérabilités, et indique quels modèles de service peuvent être affectés.

Les auteurs dans [19] ont souligné d'autres vulnérabilités qui sont communes à toutes les organisations, mais ont besoin d'être prises en considération car elles peuvent avoir un impact négatif sur la sécurité du Cloud, ainsi que sur la plateforme sous-jacente. Certaines de ces vulnérabilités sont les suivantes :

- Lacunes dans le dépistage des employés et pauvreté des pratiques d'embauche : tous les fournisseurs n'effectuent pas des vérifications des antécédents de leurs

employés ou de leurs fournisseurs. Les utilisateurs privilégiés tels que les administrateurs Cloud ont généralement un accès illimité aux données.

- Manque d'éducation sur la sécurité : les individus continuent à être le point faible dans la sécurité de l'information et ceci est vrai dans n'importe quel type d'organisation. Cependant, dans le Cloud Computing il y a un impact plus important étant donné qu'il y a plus de personnes impliquées : les fournisseurs Cloud, les fournisseurs tiers, les fournisseurs de services, les organisations et les clients finaux.

Vulnérabilités	Description	Couches impliquées
Interfaces et APIs non sécurisées	Les fournisseurs Cloud offrent des services qui peuvent être accessibles via des APIs. La sécurité du Cloud dépend de la sécurité de ces interfaces. De plus, les APIs Cloud sont encore immatures et sont donc fréquemment mises à jour. Un bug peut introduire une autre faille de sécurité dans une application. [19,21]	SPI
Vulnérabilités liées aux données	Co-localisation des données de différents propriétaires, suppression des données incomplètes, emplacement des données inconnu, données souvent traitées en texte clair.	SPI
Vulnérabilités liées à la virtualisation	- Vulnérabilités liées aux VM : isolation des VM, migration incontrôlée des VM entre des serveurs, copies incontrôlées des VM,…) - Vulnérabilité liées à l'hyperviseur : peuvent être exploitées afin de prendre le contrôle des VM ou de compromettre l'hyperviseur. - Vulnérabilités liées au Réseau virtuel : partage des ponts virtuels par plusieurs VM.	I
Vulnérabilités liées à IP	il s'agit des types d'attaques courantes comme l'attaque de homme-au-milieu, le détournement d'adresse IP, le DOS/DDoS. [21]	SPI

Tableau 2.1: Les vulnérabilités dans le Cloud Computing.

2.4.2 Les menaces

Une menace dans le Cloud est un événement potentiel, qui peut être malveillant, ou accidentel (échec d'un périphérique de stockage par exemple) compromettant les ressources Cloud. [21,19]

Le tableau 2.2 représente les principales menaces dans le Cloud Computing. Comme le tableau 2.1, il indique quels modèles de services Cloud sont exposés à ces menaces.

Les auteurs mettent d'avantage l'accent sur les menaces qui sont associés aux données et l'usage de la virtualisation.

Menaces	Description	Couches impliquées
Menaces internes	Problèmes des administrateurs Cloud Erreurs humaines ne peuvent pas être contrôlées. Selon [42], 75% des pertes donc dues à des erreurs humaines.	SPI
Détournement de comptes, ou de services (hijacking)	Un détournement de comptes, ou d'instances de service, peuvent servir de base à d'autres activités malveillantes telles que l'accès à des données sensibles, la manipulation des données et redirection d'une transaction. [21,19]	SPI
Balayage des données	Comme les données ne peuvent pas être complètement effacées à moins que l'appareil soit détruit, les attaquants peuvent être en mesure de récupérer ces données. [21]	SPI
Fuite des données	Les données confidentielles peuvent être compromises, modifiées ou supprimées.	SPI
Interruption de service	Due à des catastrophes naturelles comme des tremblements de terre, ou à une surcharge du système par à utilisateur malveillant. [42]	SPI
Manipulation des données utilisateurs	Les utilisateurs attaquent les applications web en manipulant les données envoyées par leurs composantes d'applications aux serveurs d'applications, comme injection SQL. [19]	S
Menaces liées à la virtualisation	Création, et injection de VM malveillantes dans le référentiel du fournisseur. Migration des VM à chaud expose leurs contenus au réseau. Usurpation d'adresse IP (Spoofing), et reniflement du réseau virtuel (Sniffing) Saut de VM (VM Happing) : l'accès d'une VM à une autre VM en exploitant les vulnérabilités de l'hyperviseur. [19,52] Evasion de VM (VM Escape) : l'exploitation de l'hyperviseur afin de controler l'infrastructure sous-jacente. [19,52]	I

Tableau 2.2 : Les menaces dans le Cloud Computing.

2.4.3 Les attaques

Les menaces peuvent profiter de certaines vulnérabilités pour compromettre le système. Une attaque dans le Cloud est une action visant à nuire aux ressources de celui-ci. [21]

Exemples d'attaques dans le Cloud Computing :

- Dos/ DDos

  L'attaque par déni de service a pour but de rendre indisponibles les ressources ou les services. Cette attaque consiste généralement à inonder un serveur de requêtes, afin qu'il ne soit plus en mesure de répondre aux requêtes valides. En effet, en présence d'une inondation les fournisseurs Cloud offrent plus de puissance de calcul pour servir le grand nombre de demandes (y compris les requêtes non valides) , ce qui peut causer l'indisponibilité d'un service [21,60,61,48,52] . Bien que théoriquement, le Cloud Computing offre un nombre illimité de ressources. Il dépend toujours de la façon dont les services sont configurés, et de leurs emplacements géographiques. Cette attaque est difficile à reconnaitre, car il est difficile de distinguer entre un véritable pic de la demande pour l'utilisateur du service, et une attaque DOS, car les deux créent des profils similaires quant à l'utilisation des données. [42]

  Lorsqu'un DoS (Denial of Service) est provoqué par plusieurs machines on parle alors de DDoS (Distributed Denial of Service). Dans ce cas le Cloud Computing peut atteindre un état de perte total, et ne peut satisfaire aucune requête des utilisateurs valides.

  Ce type d'attaque touche les couches infrastructure, plateforme et application, et affecte la disponibilité des services et des données. [21]

- Attaque par injection de malware

Consiste à injecter un service malicieux, ou une VM dans le système Cloud. Les malwares Cloud affectent les services Cloud en changeant (ou bloquant) les fonctionnalités Cloud [61,48,21]. Ce type d'attaque touche les couches infrastructure, plateforme et application et affecte l'intégrité des services.

- Attaque sur la virtualisation

Il existe principalement deux types d'attaques effectuées sur la virtualisation : l'évasion de machine virtuelle (VM escape) et Le détournement de l'hyperviseur (Hyperjacking):

- *VM escape*

Les VM partagent les ressources de la machine hôte. La virtualisation fournie l'isolation entre les VM et entre les VM et la machine hôte. L'évasion de VM se produit quand l'isolation entre les VM est compromise. Dans cette attaque, un programme s'exécutant sur une VM s'échappe de son encapsulation VM en brisant la couche d'isolation pour interagir directement avec l'hyperviseur, afin d'obtenir l'accès aux VM fonctionnant sur la machine physique, et même à la machine hôte. Si l'attaquant accéde à la machine hôte, il acquière les privilèges administrateurs. [21,57,58,52]

- *Hyperjacking*

Le détournement de l'hyperviseur consiste à implanter un hyperviseur malveillant, ce dernier prendra le contrôle de toute la machine hôte, et de ce fait, contrôlera toutes les interactions entre le système cible et le matériel. [36,58,59]

Ces deux attaques touchent la couche infrastructure. [21]

- Attaque de type homme au milieu (man-in-the-middle)

Cette attaque se produit si SSL (Secure Socket Layer) n'est pas correctement configuré dans le Cloud. L'attaque de l'homme au milieu a pour but d'intercepter les communications entre deux parties afin de consulter, capturer et contrôler la communication en toute transparence [21,60,61]. Ce type d'attaque touche les couches infrastructure, plateforme et application, et affecte la sécurité et la confidentialité des données. [21]

- Attaque d'hameçonnage (Phishing)

Les attaques d'hameçonnage sont bien connues pour manipuler les sites web et rediriger l'utilisateur vers un faux lien pour obtenir ses données sensibles. Dans le Cloud Computing il est possible qu'un attaquant utilise le service Cloud pour héberger un site d'hameçonnage, afin de détourner les comptes et les services (Account Hijacking) des autres utilisateurs dans le Cloud. Ce type d'attaque touche les couches infrastructure, plateforme et application, et affecte la confidentialité des informations sensibles des utilisateurs. [21]

- Usurpation d'adresse IP (IP Spoofing)

C'est une technique qui consiste à remplacer l'adresse IP de l'expéditeur d'un paquet IP par l'adresse IP d'une autre machine. Cette attaque vise le détournement de session (Account Hijacking) par attaque de déni de service (DoS) dans le but d'inonder la victime de requêtes. [35]

2.5 La sécurité des données dans le Cloud Computing

La sécurité des données et la protection de la confidentialité dans le Cloud Computing est similaire à la sécurité traditionnelle des données et la protection de la confidentialité dans les environnements traditionnels. Cependant à cause de la

caractéristique d'ouverture, et de multi-location de Cloud Computing, la sécurité et la confidentialité des données est particulière. [23]

2.5.1 La notion de vie privée

La notion de vie privée est difficile à définir. Elle a été caractérisée comme un droit dans la déclaration des droits de l'homme.

La perception de la vie privée peut changer selon la nature de l'information et son utilisateur légitime. Elle est différente dans chaque pays, culture et juridiction. [23,75]

La définition de la notion de vie privée adoptée par OECD[4] est *"Any information relating to an identified or identifiable individual"*.

Une autre définition populaire fournit par AICPA[5] et CICA[6] est *"The rights and obligations of individuals and organizations with respect to the collection, use, retention, and disclosure of personal information .*

D'une manière générale, la vie privée est associée à la collecte, l'utilisation, la communication, le stockage et destruction des données à caractère personnel. L'identification des informations privées dépend des scénarios d'application spécifiques et des lois, et c'est la tâche principale pour protéger la vie privée. [23**]**

---

[4] Organisation for Economic Co-operation and Development : est une organisation internationale d'études économiques. Son but est de promouvoir les politiques qui amélioreront le bien-être économique et social partout dans le monde.

[5] American Institute of Certified Public Accountants

[6] Canadian Institute of Chartered Accountants

2.5.2 Les informations privées

Les données utilisateur qui ont besoin d'être protégées sont les suivantes : [43,54]

- Informations personnelles

  Incluent toute information qui peut être utilisée pour identifier un individu, comme le nom, et l'adresse… Elles incluent aussi des informations qui peuvent être corrélées avec d'autres informations pour identifier ou localiser un individu, par exemple : information sur les relations sociales, code postal, adresse IP et numéro de carte de crédit.

- Informations sensibles

  Elles nécessitent une protection supplémentaire. Elles comprennent des informations sur la religion ou la race, la santé, l'appartenance syndicale ou toute autre information considérée comme privée. D'autres informations peuvent aussi être considérées comme sensibles, comme les informations personnelles financières, information sur le rendement au travail et les informations considérée comme des informations sensibles personnelles, telles que : les renseignements biométriques, ou une collection d'images de vidéo surveillance dans les lieux publics.

- Données d'usage

  Elles sont constituées des informations collectées à partir de périphérique informatique comme l'imprimante ou les habitudes des chercheurs. Elles incluent également des informations comportementales, telles que : les habitudes d'écoute des contenus numériques, les sites web récemment visités, l'historique des produits utilisés, les endroits fréquentés, ou l'interaction sociale.

- Informations permettant la localisation d'un périphérique personnel

D'autres types d'informations qui pourraient être rattachables à un dispositif d'utilisation, par exemple : l'adresse IP, l'identificateur de fréquence radio (RFID), l'identité unique d'un matériel.

2.5.3 La sécurité des données

Il existe principalement 3 types de données dans le Cloud Computing : les données en transit (les données transmises), les données au repos (les données stockées) et les données en traitement [25]. Cependant la sécurité doit être impliquée dans chaque étape du cycle de vie des données.

Le Cycle de vie des données se réfère à l'ensemble du processus de la génération à la destruction des données. Il comporte 7 étapes. [23] (cf. Figure 2.6)

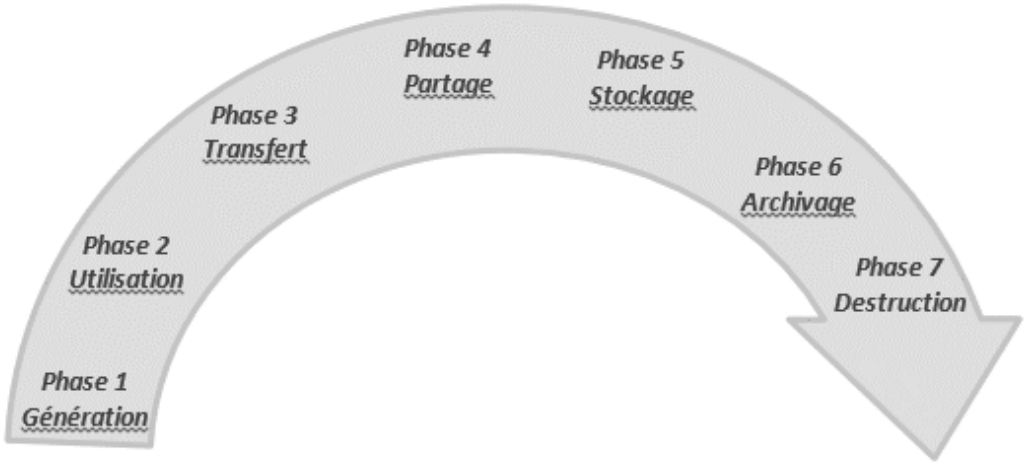

Figure 2.6: Cycle de vie des données. [23]

[23] a analysé la sécurité et la confidentialité des données associées au Cloud Computing à travers le cycle de vie des données :

- Génération (création des données)

  Dans un environnement informatique traditionnel, l'utilisateur détient, et gère ses données. Cependant, dans le Cloud Computing les données d'un utilisateur peuvent être générées par un tiers (par exemple un médecin pour les données médicales). Le problème pour le propriétaire est de garder le contrôle sur ses données crée par un autre. Pour les informations personnelles et privées, le propriétaire doit connaitre quelles informations personnelles sont collectées, et dans certains cas, d'arrêter la collecte et l'utilisation de ces informations. [23]

- Transfert

  A l'intérieur des frontières de l'entreprise, la transmission des données n'a généralement pas besoin d'être protégée, mais en dehors de ces frontières la sécurité des données transmises dans le réseau doit être assurée afin d'empêcher que les données dans ce processus ne soient interceptées, altérées, ou remplacées.

- L'utilisation

  En raison de la fonction de multi-location des modèles Cloud, les données traitées par les applications sont stockées ensembles avec les données d'autres utilisateurs. Les propriétaires de ces données ont besoin de s'assurer que l'utilisation de leurs données soit conforme aux objectifs de la collecte, et que les informations privées ne soient pas communiquées à des tiers. [39,23]

- Le partage

  Le partage des données permet l'échange d'information entre les utilisateurs. Néanmoins, les mesures de protections d'origines et les restrictions d'usage imposé par le propriétaire doivent être respectées. [23]

- Le stockage

Les données stockées dans le Cloud sont similaires à celles stockées dans les autres endroits, et ont besoin de prendre en considération les trois aspects de la sécurité de l'information : la disponibilité, l'intégrité et la confidentialité. [23]
Néanmoins, en raison de la caractéristique de co-localisation les données de différents utilisateurs sont stockées ensemble ce qui augmente les risques liés à la sécurité et au respect de la confidentialité.

- L'archivage

L'archivage des données se concentre principalement sur le support de stockage, la fourniture du stockage hors ligne et la durée du stockage. [23]

- La destruction

En raison des caractéristiques physiques des dispositifs de stockage, les données supprimées peuvent encore exister et peuvent être restaurées. Cela peut entrainer la divulgation d'informations privées [40,41,39,23]. Il n'y a aucune assurance qu'après la suppression des données par le client le fournisseur n'y ait plus accès. Aucune assurance non plus sur la suppression effective des données si le client change de fournisseur.[53]

2.6 Conclusion

La sécurité et la confidentialité des données sont les principaux problèmes dans le Cloud Computing. En effet, une fois que les données sont stockées dans le Cloud leur gestion et leur contrôle sont transférés aux fournisseurs de service. L'idée que les données soient gérées par un tiers n'est pas encore acceptée, particulièrement, pour les grandes entreprises.

Les questions relatives à la sécurité et à la confidentialité des données existent à tous les niveaux du modèles SPI (Application, Plateforme et Infrastructure), ainsi, que dans chaque étape du cycle de vie des données (création, utilisation, transfert, partage, stockage, archivage, destruction).

Les systèmes d'e-santé qui stockent les données médicales sont un des systèmes typiques qui nécessitent la confidentialité des données. L'habilité de contrôler quelles informations peuvent être révélées, et qui peut accéder à ces informations dans un tel environnement public et partagé devient une préoccupation importante.

CHAPITRE 3
PROPOSITION D'UNE ARCHITECTURE POUR LA SECURISATION DES DONNEES SUR LE CLOUD COMPUTING

3.1 Introduction

Les infrastructures médicales font face à de nombreux défis, tels que : la difficulté de partager l'information, le coût élevé des infrastructures médicales et le manque de ressources matérielles et de professionnels de santé.

L'adoption du Cloud Computing dans le domaine médical peut résoudre certaines de ces limites. En effet le Cloud Computing permet une réduction des coûts , une meilleure intégration et échange des informations médicales, une haute disponibilité, la scalabilité, la flexibilité…

Cependant, ces bénéfices ne sont pas exempts de risques, le défis majeur reste la sécurité et la confidentialité des données médicales. Les données médicales sont essentielles pour le fonctionnement de toute institution médicale. La qualité des diagnostiques et des traitements médicaux dépendent fortement de ces données. Par conséquent, celles-ci doivent être complètes, fiables, et sécurisées.

La sécurité des données médicales doit répondre aux mêmes exigences fondamentales de sécurité que n'importe quel type de données. Cependant, en raison de la nature sensible et l'importance des données médicales celles-ci nécessitent plus de vigilance.

3.2 Les exigences fondamentales de sécurité

Les systèmes médicaux utilisent des données très sensibles, la confidentialité doit donc être respectée. L'intégrité est également essentielle étant donné qu'un traitement médical incorrect basé sur des données erronées peut avoir de graves conséquences sur le patient. La disponibilité est aussi essentielle que l'intégrité car les données médicales sont nécessaires pour un traitement approprié, et doivent être accessibles continuellement.

- La confidentialité

Les données médicales sont des informations confidentielles entre un médecin et son patient. Afin que la relation patient/médecin soit efficace, il est nécessaire pour un patient de faire confiance au système médical pour protéger la confidentialité de ses données ; Si le patient sent que l'information qu'il donne à son médecin n'est pas protégée et que sa vie privée est menacée il peut être sélectif sur les informations qu'il fournit à son médecin dans le furtur. cela peut nuire à la relation patient/médecin, et entraver le diagnostic et le traitement médical approprié. [75]

En outre, les informations médicales d'un patient sont personnelles, le patient peut ne pas souhaiter que son état de santé soit révélé en dehors des professionnels de santé avec lesquels il interagit. [75] La divulgation des informations médicales concernant un patient peut avoir des conséquences négatives sur sa vie (relations familiales, relations professionnelles…). Une banque ou un employeur peuvent par exemple lui refuser un prêt ou un emploi si ses données médicales sont divulguées.

- L'intégrité

Les données médicales doivent être exactes étant donné leurs influences sur les actes médicaux. Un traitement incorrect basé sur des données erronées peut avoir de graves conséquences sur la santé des patients.

- La disponibilité

Les données doivent être disponibles tout le temps et sans interruption. Le Cloud offre généralement une haute disponibilité, cependant, un aspect important, et souvent négligé dans l'e-santé est la disponibilité des données en situation d'urgence. Une situation d'urgence est une exception à une opération normale, dans laquelle un patient - le seul qui peut rendre ses données médicales disponibles- n'est pas présent, ou n'est pas en mesure d'utiliser sa carte de santé car il est inconscient par exemple, dans ce cas ses données médicales doivent rester disponibles pour le médecin. [73]

3.3 Travaux relatifs

Il existe un certain nombre de travaux se rapportant à notre sujet. Quelques travaux ont proposé des approches pour préserver la sécurité des données médicales dans le Cloud Computing, d'autres ont analysé cette question d'une manière générale.

[66] propose une architecture de gestion de la confidentialité appelée P^3HR (Privacy-aware Patient-controlled Personal Health Record) pour le contrôle des dossiers médicaux par le patient. Les identités des patients sont cachées par le système, de telle sorte qu'il est impossible de faire le lien entre un patient et son dossier médical.

Le système P^3HR utilise des systèmes d'authentification basés sur les cartes de santé, la carte de santé d'un patient stocke ses informations personnelles sous

une forme cryptées. D'après le prototype, une implémentation basée uniquement sur un nom d'utilisateur et un mot de passe est utilisée, ce qui n'est pas suffisant.

[64] propose un mécanisme de contrôle d'accès pour assurer la confidentialité des données dans le Cloud, ce mécanisme est basé sur 2 protocoles : ABE (Attribute Based Encryption) pour assurer la confidentialité des données, et ABS (Attribute Based Signature) pour l'authentification des utilisateurs. ABE est combiné à ABS pour assurer l'anonymat des utilisateurs qui stockent leurs données dans le Cloud. La distribution des clés et des attributs est faite d'une manière décentralisée.

[72] donne un aperçu des approches communes utilisées pour préserver la confidentialité des Cloud d'e-santé, ces approches sont classées en deux catégories : les approches cryptographiques (basées sur des techniques de cryptage) et les approches non cryptographiques (basées principalement sur des contrôles d'accès).

Les auteurs ont analysé et comparé ces approches en se basant, principalement, sur le respect des exigences de sécurité. Ils ont également souligné les avantages et les inconvénients de chaque approche.

[67] traite un problème souvent négligé : la sécurité de la plateforme client. Pour répondre à ce problème les auteurs ont proposé une architecture de sécurité basée sur les TVD (Trusted Virtual Domain) dans laquelle les environnements d'exécution des applications sont partitionnés en domaines distincts isolés les uns des autres. L'objectif principal est la séparation des données médicales des autres données , pour cela, les données médicales sont gardées dans un domaine virtuel privé, et ne sont accessibles que par les utilisateurs autorisés dans ce domaine.

[74] propose l'utilisation du Cloud hybride pour le partage et l'intégration des dossiers médicaux. Les données médicales sont stockées sur le Cloud privé de l'hôpital ainsi que sur le Cloud public du fournisseur de service. Pour assurer la confidentialité, les données médicales sont cryptées.

Les dossiers médicaux peuvent être décryptés uniquement par la clé privée du patient qui est divisée aléatoirement, et stockée dans 2 endroits différents (le serveur de l'hôpital, et la carte de santé du patient). Les auteurs traitent le cas des situations d'urgence en proposant l'utilisation de clé de décryptage spéciales fournie par un tiers pour le décryptage des dossiers médicaux quand le patient n'est pas présent.

[62] propose de segmenter les données médicales des patients, et de les stocker sur trois emplacements différents (un système de stockage local, et deux plateformes Cloud commerciales) en utilisant l'algorithme RAID-3 (Redundant Array of Inexpensive/ Independent Disk).

En utilisant RAID-3, les données segmentées et stockées sur chaque site Cloud n'ont pas de sens, et ne peuvent pas être utilisées seules, ce qui améliore la sécurité et la confidentialité des patients. Afin d'assurer l'intégrité des données, une fonction de hachage MD5 (Message Digest Algorithm 5) est utilisée.

[65] propose un mécanisme dans lequel la plateforme Cloud est utilisée pour l'échange des dossiers médicaux ainsi que pour protéger la vie privée du patient.

Le respect de la vie privée du patient est atteinte avec : 1- l'anonymat du patient (en utilisant des pseudonymes), 2- le cryptage des parties sensibles des dossiers médicaux, 3- la caractéristique d'«unlinkability » dans laquelle il est impossible de faire le lien entre un patient et son dossier médical.

Pour accéder aux dossiers médicaux, les auteurs proposent l'utilisation des cartes de santé. Celles-ci sont obligatoires pour l'accès aux données, car elles contiennent les informations nécessaires pour chercher et décrypter les données médiales.

La méthode proposée prend en compte les situations d'urgence, quand la carte de santé du patient n'est pas disponible, ou quand le patient est inconscient et n'est

pas en mesure d'utiliser sa carte. Dans ce cas le médecin peut accéder au dossier médical d'un patient sans la présence obligatoire de ce dernier.

[63] propose un système sécurisé d'e-santé appelé SASHA pour assurer la sécurité des dossiers médicaux stockés dans le Cloud. Le système utilise le cryptage homomorphique pour la confidentialité des données, et un mécanisme de contrôle d'accès pour gérer les droits d'accès aux données.

[68] propose de séparer les données sensibles des données non sensibles. Seules les données non sensibles sont externalisées vers le Cloud public, les données sensibles restent dans le Cloud privé détenu et contrôlé par l'utilisateur.

Selon les auteurs, pour que le Cloud hybride soit efficace, il faut que la plus grande quantité de données traitées se fasse dans le Cloud public, sinon l'utilisation du Cloud hybride est insignifiante. Une tâche importante reste à notre avis à déterminer, et à minimiser au maximum l'ensemble des données sensibles.

[71] présente une analyse documentaire sur les défis concernant l'utilisation du Cloud Computing dans le domaine de l'e-santé , ces défis ont été classé en 6 dimensions : techniques, de confidentialité, légaux, organisationnels, économiques et médicaux.

[69] traite l'aspect de la confidentialité dans les systèmes d'e-santé en tant que problème de communication étant donné que ces systèmes sont de plus en plus distribués, et nécessitent l'interopérabilité de plusieurs sous-systèmes.

Les auteurs examinent brièvement quelques travaux traitant la confidentialité dans les systèmes d'e-santé, ces travaux ont été divisés en 2 catégories : les approches qui assurent la confidentialité du patient et celles qui assurent la confidentialité du médecin. Les auteurs soulignent la nécessité d'une approche formelle pour adresser ces problèmes.

[73] a introduit le concept des Cloud d'e-santé. Une architecture générique, son potentiel pour améliorer la qualité des soins médicaux ainsi que ses défis (classés en 2 catégories : technique et non techniques) y sont présentés.

Une analyse documentaire a été présentée concernant les recherches sur les Cloud d'e-santé, ces recherches ont été regroupées par les auteurs en 4 catégories: les solutions basées sur le stockage Cloud, les solutions basées sur les plateformes, les modèles d'implémentation des Cloud d'e-santé, et les solutions pour la sécurisation des systèmes Cloud d'e-santé.

[70] présente quelques aspects concernant les problèmes de sécurité et de confidentialité des Cloud d'e-santé, ces problèmes sont classés en 2 types par les auteurs : techniques, et légaux.

Parmi les travaux présentés, 5 traitent de notre problématique. Une comparaison des solutions proposées est représentée dans le tableau 3.1. L'objectif de ce tableau est de mettre en évidence les points qui nous semblent importants pour notre travail, à savoir le respect des exigences fondamentales de sécurité, et la méthode utilisée pour réaliser chacune d'elle.

Art.	Objectif	Respect des exigences de sécurité			Méthode utilisée pour assurer la Confidentialité	Méthode utilisée pour assurer l'intégrité	Méthode utilisée pour gérer les situations d'urgence
		C	I	D			
62	Stockage distribué des dossiers médicaux	Oui	Oui	Non	Cryptage et segmentation des données, et stockage de chaque segment sur un emplacement cloud différent	Fonction de hachage MD5	/
63	Sécurisation des dossiers médicaux dans le Cloud	Oui	Non	Non	Cryptage homomorphique, et contrôle d'accès basé sur les attributs	/	/
67	Sécurisation de la plateforme de l'utilisateur final durant l'accès aux données médicales dans le cloud	Non	Non spécifiée	Non	Construction d'un domaine privé pour les données médicales qui sont cryptées avec une clé accessible uniquement dans ce domaine	/	/
74	Sécurisation du partage des dossiers médicaux dans le Cloud hybride	Non	Oui	Oui	Cryptage des données médicales, Segmentation et distribution des clés de décryptage sur 2 emplacements	Signature numérique	Création d'une clé d'urgence de décryptage par un tiers en utilisant les informations fournies par le médecin
65	L'échange et le partage sécurisé des dossiers médicaux dans le Cloud	Oui	Oui	Oui	Anonymat des patients, Cryptage des parties sensibles des dossiers médicaux, Contrôle d'accès, Authentification des utilisateurs	Signature numérique	Récupération des informations nécessaires pour l'accès aux données de 2 parties indépendantes

Tableau 3.1 : *Comparaison des approches proposant une solution pour sécuriser les données médicales dans le Cloud*

3.4 Analyse et critique des travaux relatifs

Nous observons que la majorité des solutions, hormis les travaux [67] et [74], remplissent l'exigence de confidentialité. En effet, [67] ne protège pas contre les menaces internes provenant du même TVD, la méthode proposée protège uniquement des menaces externes au TVD. Dans [74], l'accès aux données dans les situations d'urgence n'est pas assez sûr car les utilisateurs voulant accéder aux données médicales dans ce cas ne sont pas authentifiés, et peuvent accéder aux données en présentant uniquement l'ID du patient, l'ID du dossier médical et l'identité de l'hôpital ce qui n'est pas suffisant. Ces informations peuvent etre connues par beaucoup d'utilisateurs travaillant dans le même hôpital par exemple.

Pour assurer la confidentialité, la plupart des travaux utilisent des techniques de cryptage pour protéger les données et des mécanismes de contrôles d'accès pour limiter les accès aux personnes non autorisées. Cependant ces solutions ont des limitations :

Le cryptage introduit des calculs lourds (distribution des clés, cryptage et décryptage des données,…) en plus du problème de gestion des clés. En effet, la seule tâche de décider qui est responsable de la gestion des clés est difficile. Idéalement, c'est le propriétaire des données. Mais à l'heure actuelle, parce que les utilisateurs n'ont pas suffisamment d'expertise pour gérer les clés, ils confient, généralement, leur gestion aux fournisseurs. Comme ces derniers ont besoin de maintenir les clés de plusieurs utilisateurs, la gestion des clés devient plus complexe. La gestion des clés par les fournisseurs soulève, également, des problèmes de confidentialité étant donné qu'ils peuvent déchiffrer les données cryptées.

Le problème avec le contrôle d'accès est qu'un contrôle d'accès strict améliore la sécurité mais limite la gestion des situations d'urgence.

L’objectif principal de la confidentialité des données est de préserver la vie privée du propriétaire. Cependant dans le domaine médical, le seul fait de savoir qu’une personne possède un dossier médical implique qu’elle reçoit des soins ce qui représente une violation de sa vie privée, même si ses données médicales restent confidentielles. C’est le cas dans [63] dans lequel les patients sont enregistrés par les administrateurs qui entrent leurs données personnelles dans le système. Ainsi les administrateurs connaissent l’identité du patient même si ils n’ont pas accès à leurs données médicales. Ce qui pose un problème de confidentialité.

Une propriété importante et souvent non considérée dans l’e-santé est l’anonymat des patients. L’anonymat consiste à cacher les identités des patients généralement avec des techniques de « pseudonymisation » consistant à remplacer les informations personnelles permettant d’identifier un patient (nom, tel. Adresse…) par des pseudonymes.

Selon [66] la divulgation de certaines informations à une personne non autorisée ne signifie pas nécessairement la perte de la confidentialité/vie privée, si la personne non autorisée n’est pas en mesure de relier les informations divulguée à un individu spécifique. Cette caractéristique est appelée « unlinkability ».

L’avantage de ces deux caractéristiques est qu’elles résolvent le problème des menaces internes. Ces menaces proviennent des employés internes du fournisseur Cloud (administrateurs Cloud). Ce type de menace est difficile à sécuriser du fait que les administrateurs Cloud sont des utilisateurs légitimes, et ne peuvent être détectés par le système de sécurité. L’avantage avec l’anonymat est que même si un administrateur accède à un dossier médical, il ne peut deviner à qui il appartient. De plus les dossiers médicaux anonymes peuvent servir à la recherche sans dévoiler les identités des patients.

Cependant le problème posé par l’anonymat est l’authentification des utilisateurs. Comment vérifier l’authenticité d'un utilisateur sans connaitre son identité ? Les auteurs dans [64] proposent l’utilisation des signatures numériques

telle que : ABS (Attribute-Based Encryption) qui permet d'identifier l'utilisateur comme une personne valide sans révéler son identité, ce qui permet en plus de vérifier l'authenticité du message.

L'intégrité des données est aussi importante que la confidentialité étant donné que les traitements et les diagnostics médicaux dépendent de l'exactitude des données. La plupart des travaux, hormis [63], répondent à cette exigence. La méthode la plus utilisée pour assurer l'intégrité des données est la signature numérique. Un processus important est la vérification de la signature, pour s'assurer que les données contenues dans un dossier médical n'ont pas été modifié ou altéré et qu'elles correspondent à celles créées par le médecin.

La disponibilité en situation d'urgence se pose quand la présence du patient est obligatoire pour l'accès aux données. Ce cas n'est traité que par deux travaux [74] et [65] qui proposent la récupération des informations nécessaires pour l'accès aux données (génréralement la clé de décryptage) à partir d'un tiers de confiance. La gestion de cette situation soulève de nombreuses questions, telles que : Qui doit être responsable de cette tâche ? Dans quel cas cette personne a la permission d'ouvrir le dossier ? Doit-on impliquer des actions légales ?

Un autre problème souvent soulevé est l'interopérabilité entre les systèmes et les données. Le problème se pose quand les services d'e-santé sont fournis à partir de plusieurs emplacements Cloud, ou à partir d'un emplacement local et un emplacement Cloud comme c'est le cas pour [62] et [74]. Chaque fournisseur utilise ses propres protocole, et formats de données. Ce qui rend difficile l'intégration et la communication des informations médicales entre les différentes institutions.

La majorité des travaux ne respectent pas toutes les exigences fondamentales de sécurité à part [65]. Néanmoins ce dernier soulève un problème de portabilité du mécanisme étant donné qu'il repose sur NHI (National Health Insurance) et HCA (Health Cenrtification Autority) pour l'enregistrement des patients et des médecins, l'accès aux données dans les situations d'urgence…

Selon [74], la confidentialité des données médicales ne concernent pas uniquement les questions de sécurité/vie privée mais également les questions juridiques.

L'aspect légal joue un rôle crucial dans l'adoption du Cloud Computing par le secteur de la santé. Les données médicales sont des informations confidentielles entre un patient et son médecin. L'utilisation de telles informations sensibles nécessite un cadre légal et éthique qui doit protéger la vie privée des patients, et sécuriser leurs données médicales.

3.5 Proposition d'une solution pour sécuriser les données médicales dans le Cloud Computing

Notre contribution se base sur le système proposé par [66] que nous tenterons d'améliorer. Nous indiquerons tout d'abord, nos motivations pour le choix de ce système. Nous décrirons ensuite, l'architécture et le fonctionnement du système proposé par [66]. Nous finirons cette partie par la description de notre proposition.

3.5.1 Motivations

Notre choix s'est porté sur ce travail pour les raisons suivantes :

- le système est simple à comprendre et à utiliser.
- le système n'utilise pas le cryptage pour assurer la confidentialité des données. Il se base sur l'anonymat, ce qui évite le cryptage d'une masse importante de données médicales, et les problèmes de gestion des clés (et les questions de confidentialité qu'ils soulèvent).
- étant donné que les dossiers médicaux sont stockés sous forme anonyme, le problème des menaces internes est résolu.

3.5.2 Présentation du système P^{3}HR

[66] propose une architecture de gestion de la confidentialité appelée P^{3}HR (Privacy-aware Patient-controlled Personal Health Record) pour le contrôle des dossiers médicaux par le patient. Afin de préserver la vie privée des patients, leurs identités sont cachées par le système, de sorte qu'un dossier particulier ne puisse être associé à un individu spécifique.

3.5.2.1 Architecture du système P^{3}HR

L'idée principale sur laquelle repose l'architecture du système P^{3}HR est que les identités des patients sont rendus anonymes en supprimant leurs identifiants, et en les remplaçant par des pseudonymes. Ainsi même si une personne non autorisée obtient l'accès aux dossiers médicaux, elle ne peut pas associer un dossier médical à un patient particulier. L'architecture du système est représentée dans la figure 3.1.

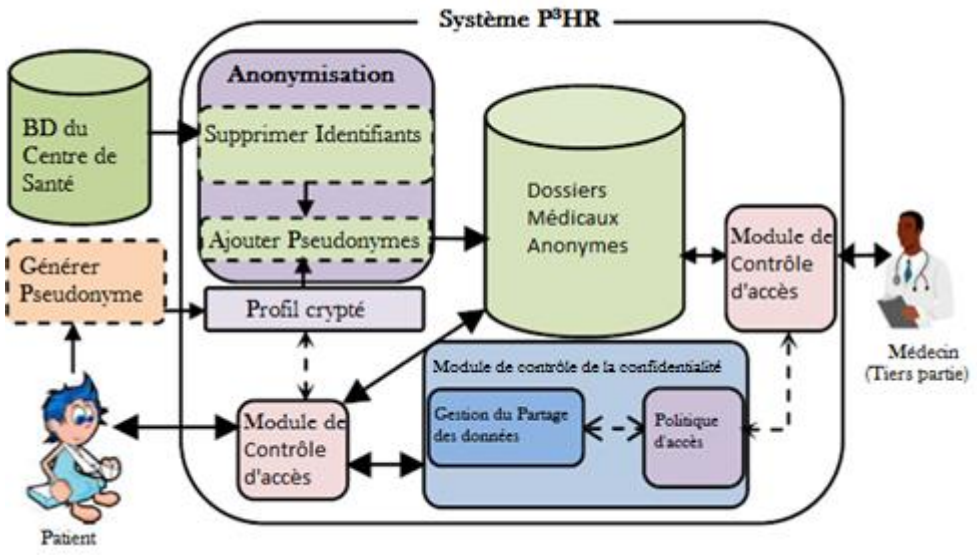

Figure 3.1: Architecture du Système P^{3}HR.

- Module d'Anonymisation

Le module d'anonymisation supprime tous les identifiants des dossiers médicaux afin de dissocier les liens entre les patients et leurs dossiers médicaux. Afin de permettre à une personne autorisée d'accéder au dossier médical d'un patient, le système a besoin d'associer chaque dossier au patient correspondant. Pour cela, le patient crée son ID unique (pseudonyme). Le pseudonyme d'un patient est ajouté à tous ses dossiers médicaux durant le processus d'ajout des dossiers. Ainsi, un dossier dans la base de données contient le pseudonyme du patient avec ses informations médicales. Le pseudonyme est crypté et stocké sur la carte de santé du patient dans son Profil. La figure 3.2 décrit le fonctionnement du module d'anonymisation.

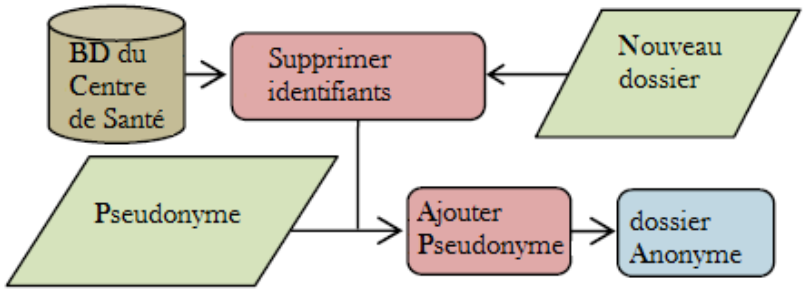

Figure 3.2: Opération d'anonymisation.

- Profil du patient

Le profil du patient contient ses informations personnelles identifiables, comme : le nom, l'adresse, le Tel, le numéro de sécurité sociale…il contient également d'autres informations privées telles que le pseudonyme. Le profil est crypté avant d'être stocké sur la carte de santé. (cf. Figure 3.3)

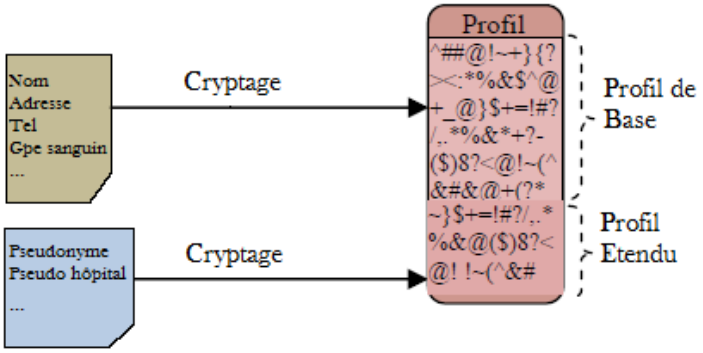

Figure 3.3: Profil du Patient.

- Module de Contrôle d'Accès

Les dossiers médicaux sont accessibles par les patients et les professionnels de santé. Le système propose deux modules de contrôle d'accès : un pour le patient et un pour le professionnel de santé.

Contrôle d'accès pour le patient

Un patient peut voir uniquement les enregistrements contenant son pseudonyme. Le pseudonyme doit être décrypté avec sa clé privée (détenu uniquement par lui).

Le pseudonyme est obligatoire pour l'accès aux données. Seul un patient légitime possède la clé privée de décryptage et peut décrypter son pseudonyme. Le module de contrôle d'accès du patient gère :

- l'accès aux données.
- le partage des données.
- la politique d'accès aux données.

Contrôle d'accès pour le professionnel de santé

Les professionnels de santé ont des droits limités, ils peuvent avoir seulement accès aux dossiers médicaux dont ils ont l'autorisation. Les identités des professionnels de santé sont vérifiées grâce à des moyens externes (lecteurs de cartes).

Authentification

L'authentification est basée sur des cartes de santé (cf. Figure 3.4). Pour cela les patients et les professionnels de santé ont des cartes d'accès. Chaque patient possède une carte personnelle qui stocke les informations du Profil sous une forme cryptée. Les professionnels de santé ont des cartes professionnelles pour l'authentification.

Les lecteurs de cartes (installés dans les centres de santé) peuvent décrypter et lire les informations à partir des cartes. L'accès aux données par le professionnel de santé doit se faire en présence du patient en utilisant sa carte personnelle. Le lecteur de carte vérifie l'identité du professionnel de santé avant de lire le contenu de la carte du patient qui doit entrer son code PIN. Les cartes médicales du patient et du professionnel de santé sont donc obligatoires pour l'accès aux données.

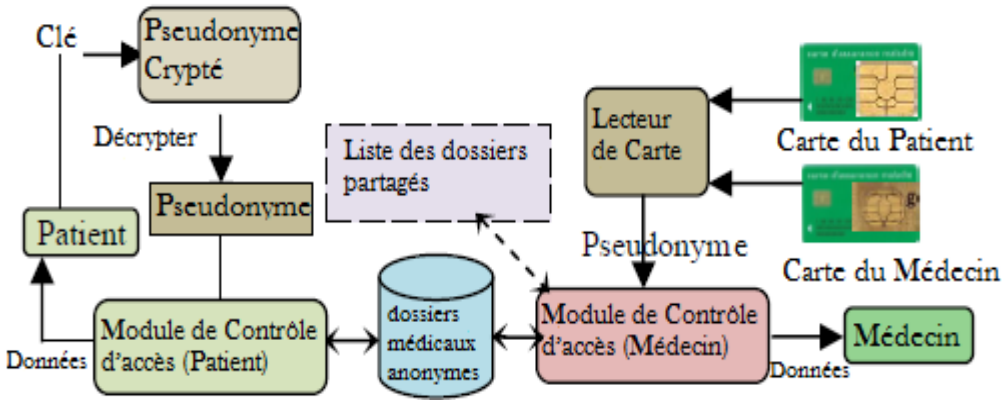

Figure 3.4: Module de Contrôle d'accès.

➢ Module de Contrôle de la Confidentialité

Chaque patient est autorisé à créer et gérer les politiques d'accès à ses données médicales par des tiers. Dans son Profil, le patient détermine les règles d'accès pour les professionnels de santé. (cf. Figure 3.5). Le module de Contrôle de la Confidentialité comprend deux éléments :

- ***politique d'accès*** : permet au patient de choisir les dossiers médicaux qui peuvent être accessibles.

- ***gestion des données partagées*** : permet au patient de gérer le partage de ses informations médicales. Il peut choisir les professionnels de santé qui peuvent accéder à son dossier médical, la durée de l'accès et les parties du dossier accessibles.

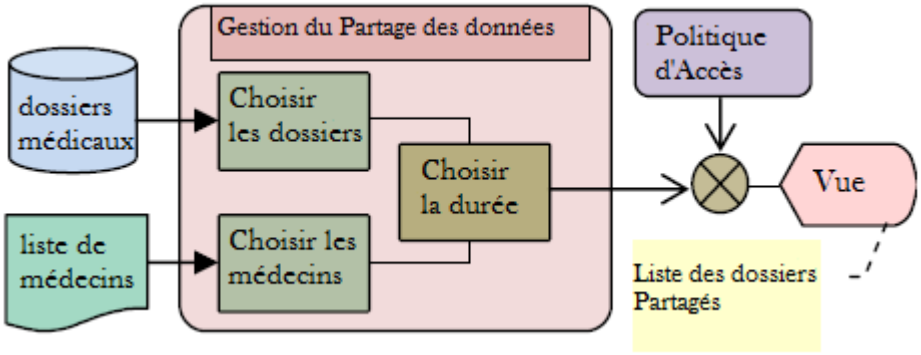

Figure 3.5: Module de contrôle de la Confidentialité.

3.5.2.2 Fonctionnement du système P^{3}HR

Quand un professionnel de santé veut accéder au dossier médical d'un patient il utilise sa carte professionnelle pour être authentifié par le système ainsi que la carte personnelle du patient. Le système récupère le pseudonyme crypté et la clé privée (pour le décryptage du pseudonyme) du patient à partir de sa carte

personnelle. Le système utilise le pseudonyme pour chercher le dossier médical correspondant dans la BD des dossiers anonymes et l'affiche au professionnel de santé sans lui dévoiler le pseudonyme. Ce dernier doit rester secret même pour le professionnel de santé.

Le patient peut accéder à son dossier médical soit de l'hôpital en utilisant sa carte personnelle, soit en utilisant un navigateur web.

Le problème qui se pose avec ce système est que la présence du patient est obligatoire pour l'accès aux données. Il est possible que le patient ne puisse pas être présent parce qu'il est inconscient par exemple, ses données doivent rester accessibles par le professionnel de santé. La gestion de cette situation n'est pas prise en charge par le système P^3HR. De plus l'intégrité des données n'est pas vérifiée. Etant donné l'importance de ce type de données pour les décisions médicales, le médecin doit être sûr de l'exactitude des données avant de les utiliser.

3.5.3 Contribution

Notre contribution vise à améliorer le travail [66]. Elle consistera principalement en:

- l'adaptation du système dans le Cloud : Centralisation du système dans le Cloud au lieu de l'installer sur le serveur de chaque hôpital. Ainsi, le système devient accessible de n'importe quel hôpital et à tout moment.
- la proposition d'une solution pour la gestion des situations d'urgence.
- la vérification de l'intégrité des dossiers médicaux.

3.5.3.1 Changements apportés au system P^3HR

- Développement du système en tant qu'application Cloud.

- Le dossier médical d'un patient contient deux types d'informations : des informations personnelles (comme le nom, l'adresse, le tel…) et des informations médicales (comme le groupe sanguin et la pathologie). Nous considérons que le propriétaire des informations médicales est celui qui les a créées à savoir l'hôpital. Ce dernier a un contrôle total sur ces informations, il peut les créer, les consulter, les modifier les supprimer, et les partager avec d'autres hôpitaux. Toutefois, les informations personnelles sont la propriété du patient qui a droit à sa vie privée. La vie privée du patient sera protégée avec l'anonymat.
- les patients peuvent uniquement consulter leurs dossiers médicaux (ils n'ont pas d'autres contrôles sur les dossiers).
- la création des pseudonymes se fait dans l'hôpital.
- tous les accès, créations et modifications de dossiers médicaux doivent être signés.

3.5.3.2 Idée générale du système proposé

La carte de santé d'un patient stocke les informations nécessaires pour l'accès à son dossier médical, à savoir, son pseudonyme crypté et sa clé de décryptage.

En situation normale, quand la carte de santé du patient est disponible, nous utilisons la même architecture proposée par les auteurs, dans laquelle le système récupère le pseudonyme crypté et la clé privée du patient à partir de sa carte personnelle.

Dans le cas d'une situation d'urgence, quand la carte de santé du patient n'est pas disponible, nous proposons de stocker les informations nécessaires pour l'accès aux données sur le Cloud. Le système récupère ces informations à partir du Cloud au lieu de les chercher sur la carte du patient. Etant donné que le Cloud n'est pas considéré comme un environnement sûr, et que les pseudonymes sont des informations privées qui doivent rester secrets, ces derniers doivent être cryptés. Pour améliorer la sécurité, la gestion des clés doit se faire dans un environnement

sûr (à savoir l'hôpital). Les clés de décryptage seront donc, en plus d'être stockées sur les cartes de santés des patients, stockées sur le serveur de l'hôpital, ainsi, nous sécurisons les accès aux dossiers médicaux en situations d'urgence, étant donné que seules les personnes autorisées (celles qui possèdent les clés de décryptage) peuvent décrypter les pseudonymes et accéder aux données.

3.5.3.3 Architecture du système proposé

Le principe sur lequel s'est basée notre architecture est le suivant : nous avons séparé les données contenues dans les dossiers médicaux en deux ensembles :

- les données sensibles : constituées des informations personnelles des patients, qui permettent de les identifier telles que : le nom, l'adresse, le tel…Pour protéger la vie privée des patients ces données seront supprimées et remplacées par des pseudonymes.

- Les données non sensibles : constituées des informations médicales des patients, qui sont inutiles sans les identités des propriétaires (patients).
  L'architecture générale de notre solution est représentée dans la figure 3.6. Notre solution est composée de deux parties principales : l'infrastructure de l'hôpital (environnement sûr) et l'infrastructure Cloud (environnement non sûr).

➢ *L'infrastructure de l'hôpital* est responsable de la création des pseudonymes et des dossiers anonymes. Tout le mécanisme de confidentialité de notre solution repose sur les pseudonymes. Le mécanisme de création des pseudonymes est donc très important ; C'est pour cette raison que la création des pseudonymes doit se faire dans un environnement sécurisé (ici l'hôpital). L'objectif principal de cette partie de l'architecture est la création des dossiers médicaux anonymes ; une fois ces derniers créés, ils sont stockés sur le Cloud.

- *L'infrastructure Cloud* permet l'accès aux dossiers médicaux anonymes. Les dossiers anonymes sont accessibles par les patients et les professionnels de santé après leurs authentifications par le système.

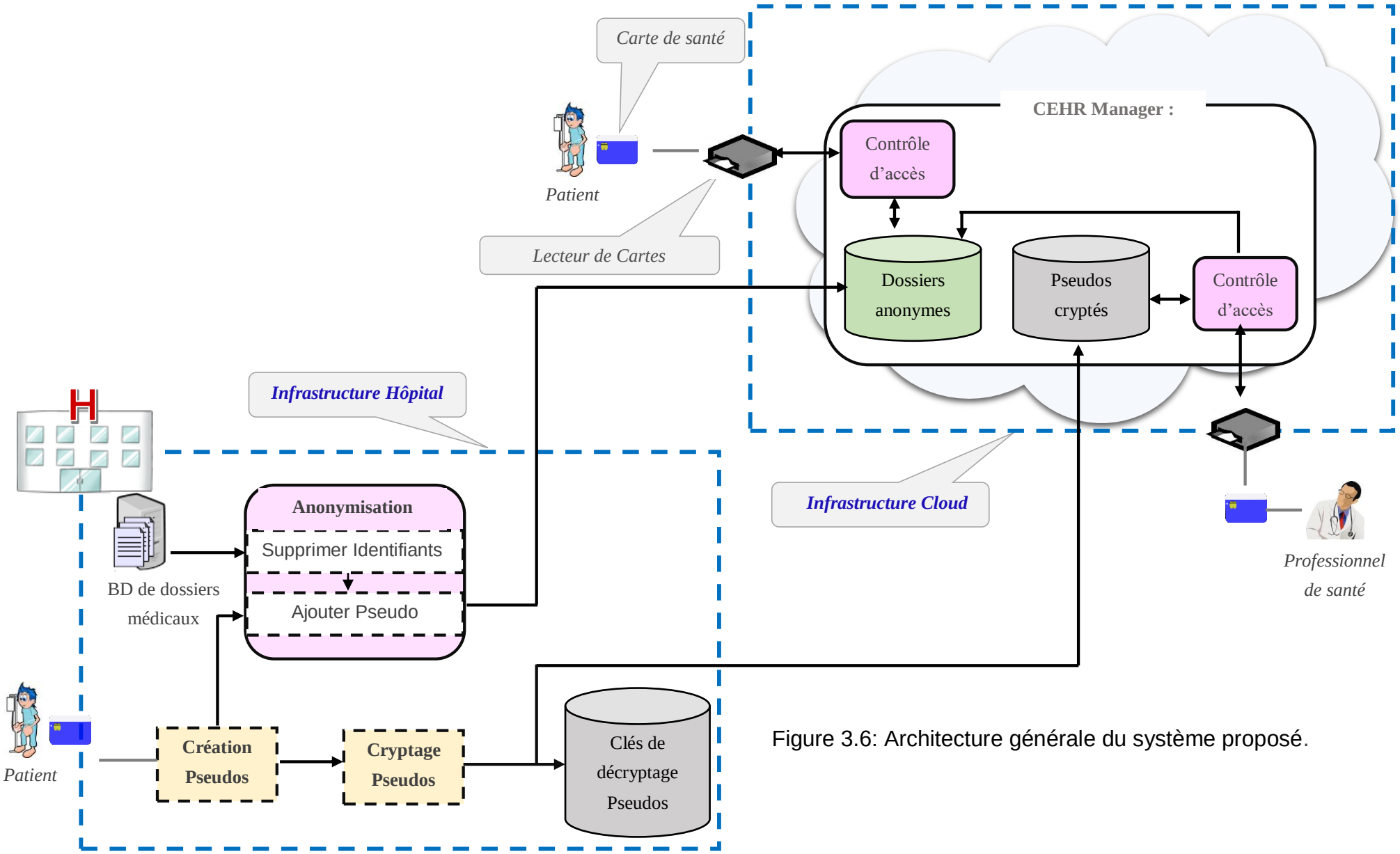

Figure 3.6: Architecture générale du système proposé.

3.5.3.4 Fonctionnement du système proposé

- *Ajout d'un nouveau dossier médical anonyme*

Les dossiers médicaux ne peuvent être ajoutés que par les professionnels de santé après leurs authentifications par le système. Ce dernier supprime les informations personnelles du patient (identifiants). Le pseudonyme du patient est récupéré, soit à partir de sa carte personnelle si elle est disponible, soit du cloud dans le cas contraire. Le pseudonyme est décrypté par le système et ajouté à toutes les informations médicales du patient. Une empreinte est créée pour le dossier médical anonyme qui doit être signé par le médecin avant d'être stocké dans la base de données des dossiers anonymes sur le Cloud. Le diagramme représentant l'ajout d'un nouveau dossier est représenté dans la figure suivante :

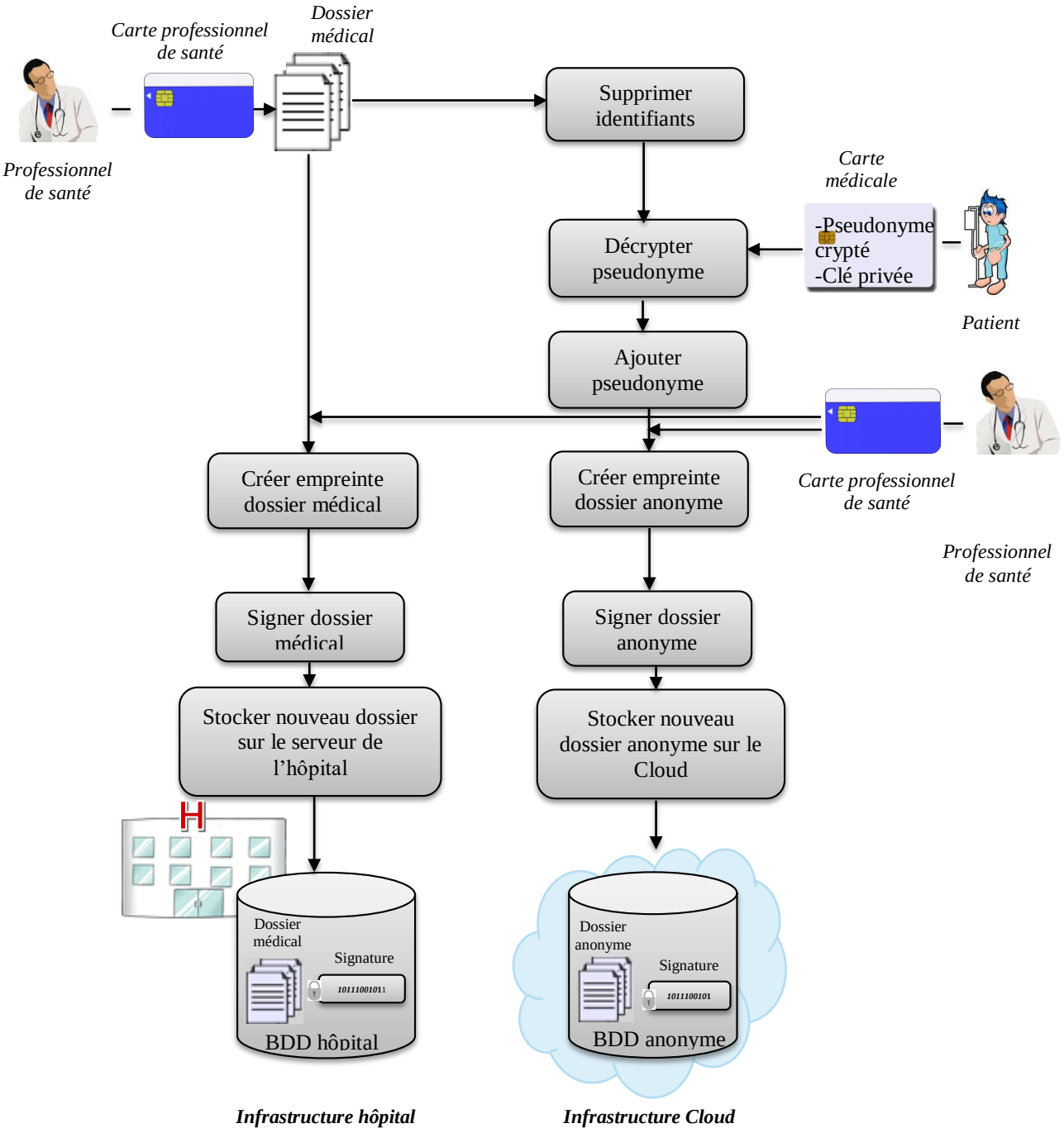

Figure 3.7 : Ajout d'un nouveau dossier médical.

- *Création des pseudonymes*

La création des pseudonymes est un mécanisme très important étant donné que la confidentialité des dossiers médicaux dépend de ce dernier. Pour créer les pseudonymes des patients, nous nous sommes basés sur le mécanisme *Unique user-generated digital Pseudonyms* proposé par [76] que nous avons modifiés. Ce mécanisme permet la création de pseudonymes uniques et aléatoires en utilisant une fonction de cryptage asymétrique RSA[7] .

Les pseudonymes sont des identifiants de sujets, en l'occurrence des patients. D'un point de vue technique, un pseudonyme est une chaine de bits ayant les caractéristiques principales suivantes :

- Ce sont des identificateurs uniques et sont utilisés pour authentifier les propriétaires.
- Ils doivent être créés aléatoirement.
- Il n'y a aucun moyen de faire le lien entre un pseudonyme et son propriétaire.

Pour répondre à la première exigence, le pseudonyme du patient sera créé à partir de son numéro de sécurité sociale ; Etant donné que les numéros de sécurité sociale sont uniques au niveau national, tous les pseudonymes seront différents.

Concernant la deuxième, et la troisième exigence, les pseudonymes seront créés avec un algorithme de cryptage RSA. RSA est basé sur la génération d'une paire de clé aléatoire et une fonction à sens unique, une fois appliquée à un message, il est extrêmement difficile de retrouver le message original. Les

[7] Abréviation tirée des trois noms de ses auteurs : R.Rivest, A.Shamir et L.Adleman

pseudonymes des patients sont utilisés pour chercher leurs dossiers médicaux, ils sont donc nécessaires pour l'accès aux données.

- L'algorithme RSA

RSA est un algorithme de cryptage asymétrique reposant sur l'utilisation d'une clé publique (qui est diffusée) pour le cryptage et d'une clé privée (gardée secrète) pour le décryptage.

Avec ce système, tout le monde peut crypter un message en utilisant la clé publique (connu par tous). Mais seul le propriétaire du message (celui qui l'a crypté) peut le décrypter avec sa clé privée (connu uniquement par lui).

m : message en clair ; *c* : message crypté ; mod : opération modulo (le reste de la division).

- **Génération des clés**
  1. Génération de deux grands nombres premiers aléatoires *p* et *q*. Calculer $n=p*q$.
  2. Clé publique : prendre un nombre *e* premier avec $(p-1)*(q-1)$. La clé publique est (e,n).
  3. Clé privée : calculer *d*, tel que : $e*d \bmod (p-1)*(q-1) \equiv 1$. La clé privée est (d,n).
- **Cryptage du message *m***

  $c = m^e \bmod n$
- **Décryptage du message *c***

  $m = c^d \bmod n$

Algorithme 3.1 : L'algorithme RSA. [77]

- Génération des pseudonymes

Le pseudonyme P d'un utilisateur ayant l'identifiant unique *numéro de sécurité sociale (NSS)* est généré en utilisant une fonction de cryptage asymétrique RSA comme suit :

- le module *n*, l'exposant public *e* et l'exposant privé *d* sont générées dans un environnement sécurisé (hôpital).

- le *nss* est concaténé avec quelques données additionnelles *data* et est crypté avec la clé publique (*e*,*n*).

- le résultat du cryptage est concaténé avec la clé publique, et forme un pseudonyme unique et aléatoire. (cf. Algorithme 2)

L'algorithme de génération du pseudonyme est le suivant :

Entrée : *NSS, data*

Sortie : P

(1) générer deux nombres premiers *p, q*. calculer $n=p*q$

(2) générer une clé publique aléatoire *e*, avec *e* premier avec $(p-1)*(q-1)$

(3) calculer la clé privée *d*, tel que : $e*d \bmod (p-1)*(q-1) = 1$

(4) générer le pseudonyme $P = E_e(NSS \| data) \| e \| n$

(5) retourner *P*

Algorithme 3.2 : La génération d'un pseudonyme unique. [76]

➢ *Accès à un dossier médical*

Les dossiers médicaux peuvent être accessibles par les patients et par les professionnels de santé.

Chaque patient peut accéder à son dossier en utilisant sa carte de santé ; Cette dernière contient son pseudonyme crypté la clé privée qui permet le décrypter, donc personne à part lui ne peut accéder à ses informations médicales.

Le professionnel de santé peut accéder au dossier médical de son patient en utilisant sa carte professionnelle pour être authentifié ainsi que la carte

personnelle du patient qui contient son pseudonyme. Le pseudonyme du patient est récupéré à partir de sa carte de santé, et est utilisé pour chercher le dossier médical correspondant dans la base de données des dossiers anonymes sur le Cloud. Quand la carte du patient n'est pas disponible, le professionnel de santé récupère le pseudonyme crypté à partir du Cloud et la clé de décryptage à partir de l'hôpital.

Le diagramme décrivant l'accès à un dossier médical par le professionnel de santé est représenté dans la figure suivante :

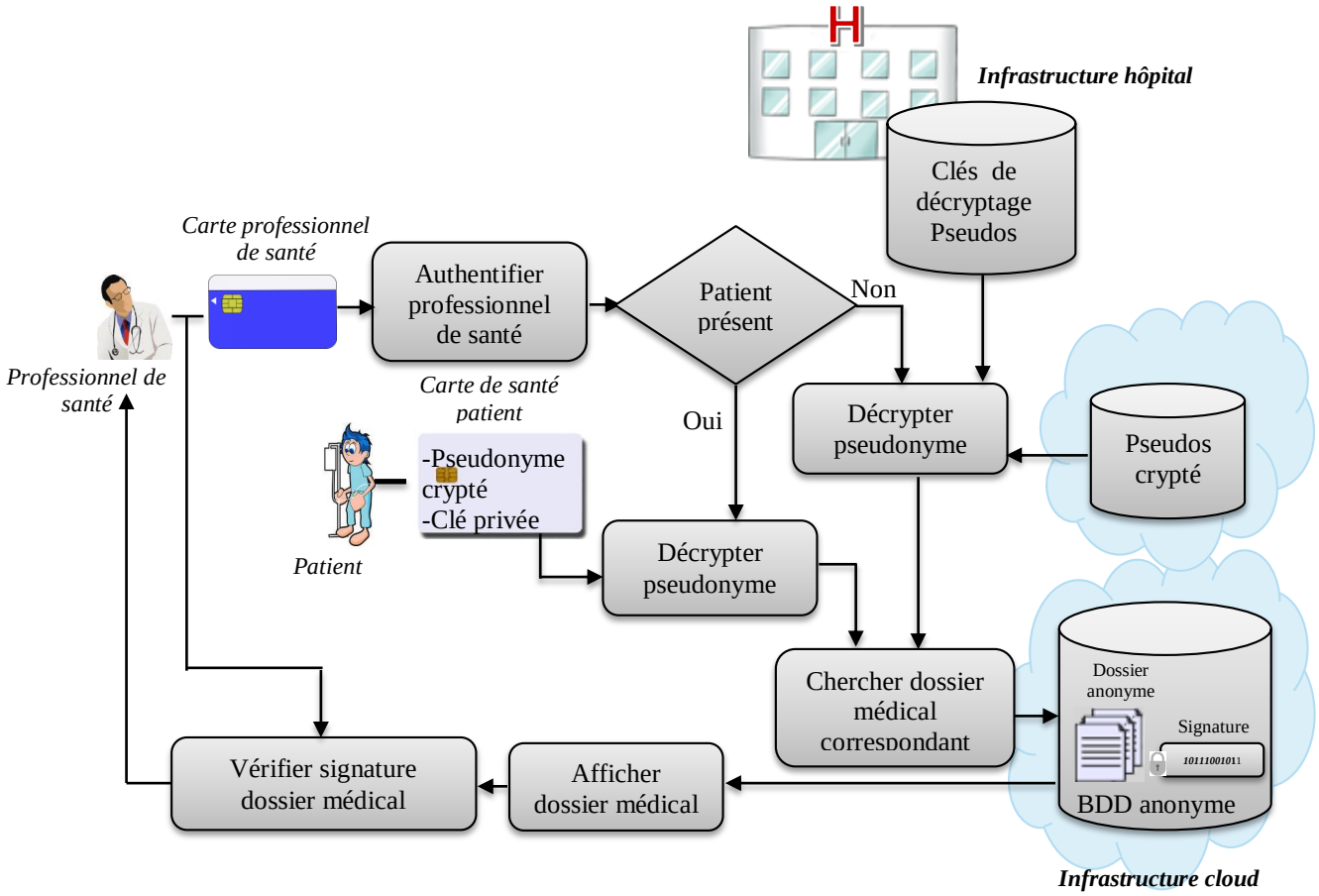

Figure 3.8: Accès à un dossier médical par le professionnel de santé.

- *Création d'une l'empreinte et la signature d'un dossier médical*

Les dossiers médicaux doivent être signés par les professionnels de santé lors de leurs créations, et après chaque modification. L'algorithme RSA peut également être utilisé pour créer et vérifier les signatures numériques. Cependant, contrairement au cryptage, la clé privée sert à crypter le message et la clé publique à le décrypter. Ainsi seul le créateur du dossier, à savoir le professionnel de santé, peut signer le dossier (avec sa clé privée : connue uniquement par lui) et tout le monde peut vérifier son intégrité (avec la clé publique : connue par tous). La création d'une signature numérique se fait comme suit :

- le professionnel de santé doit, tout d'abord, créer une empreinte (hash) du dossier médical en utilisant une fonction de hachage (nous proposons l'utilisation de la fonction de hachage SHA-256).
- le professionnel de santé crypte ensuite l'empreinte du dossier en utilisant sa clé privée stockée sur sa carte professionnelle en utilisant l'algorithme RSA.
- l'empreinte cryptée est stockée avec son dossier médical correspondant dans la base de données anonyme sur le Cloud ainsi que dans la base de données de l'hôpital.
- si une personne autorisée veut consulter un dossier médical, elle peut vérifier si les informations contenues dans ce dossier sont correctes, et qu'elles correspondent à celles crées par le professionnel de santé. Pour cela, elle utilise la clé publique pour décrypter l'empreinte et la comparer avec l'empreinte réelle du dossier. La signature numérique permet également de vérifier l'authenticité du dossier.

Les figures 3.9 et 3.10 décrivent respectivement la création de la signature et la vérification de la signature.

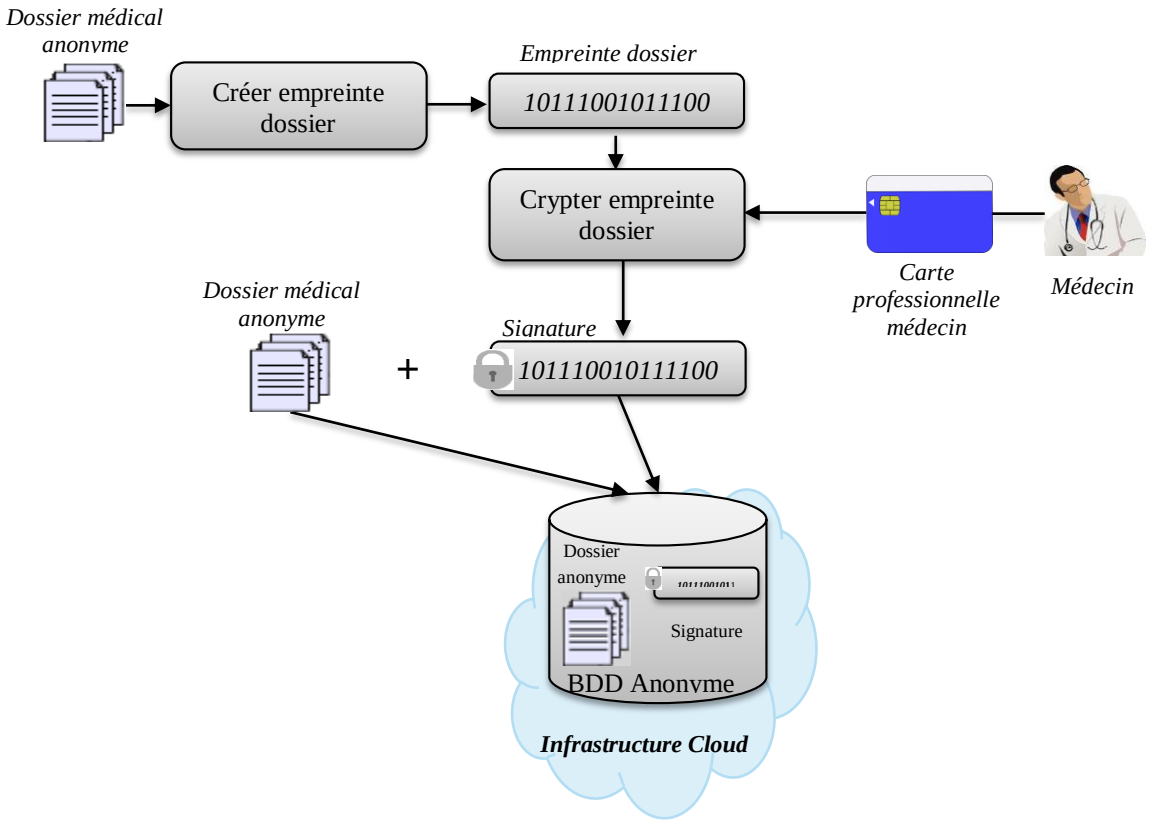

Figure 3.9: Création de la signature numérique d'un dossier médical par le professionnel de santé.

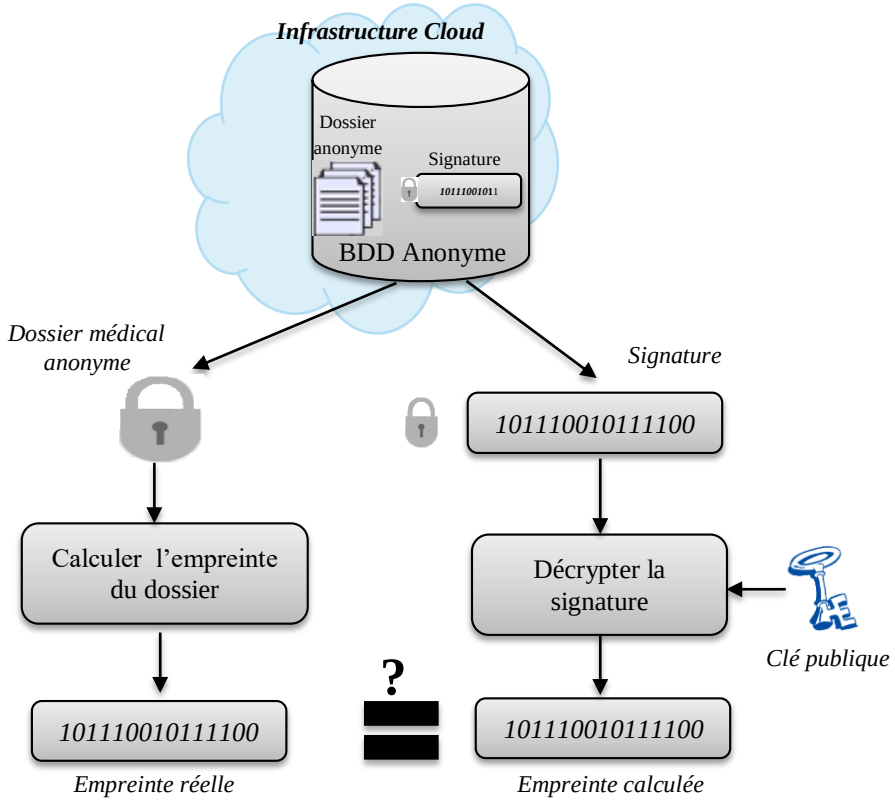

Figure 3.10: Vérification de la signature numérique d'un dossier médical.

- *Choix de la fonction de hachage*

Une fonction de hachage est une fonction qui transforme un message de taille quelconque en une chaine de bits de taille fixe ; Le résultat obtenu est appelé empreinte du message (« hash ») (cf. Figure 3.11). L'objectif de l'empreinte est de représenter les données de façon certaine tout en réduisant la taille qui sera chiffrée. L'empreinte doit être la plus courte possible et garder sa propriété principale, à savoir, l'unicité des empreintes créées. Le problème qui se pose ici est que plus l'empreinte est petite, plus il y a de chance de produire la même empreinte pour deux données différentes.

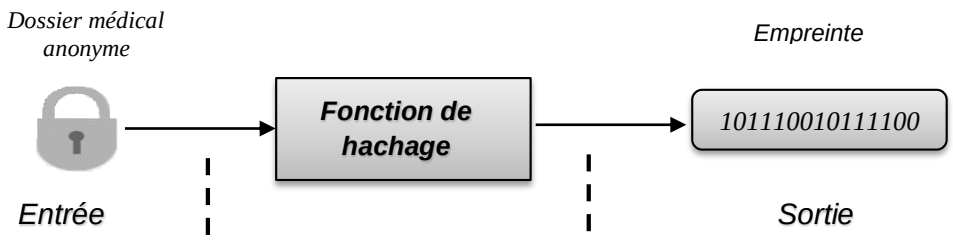

Figure 3.11: Fonction de hachage: Principe général.

Nous avons choisi la fonction de hachage SHA-256 car elle répond aux exigences de sécurité principales des fonctions de hachage, à savoir :

- fonction de hachage à sens unique : difficulté de trouver un message à partir de son empreinte.
- résistante aux collisions : difficulté de trouver deux message différents ayant la même empreinte avec SHA-256.

SHA-256 est actuellement la nouvelle norme recommandée par les autorités de certification.

3.5.3.5 Discussion/ Analyse

Nous nous somme basés dans notre contribution sur le système P^3HR proposé dans [66] que nous avons tenté d'amélioré. Notre amélioration s'est portée sur les points suivants :

- P^3HR consiste en une application locale qui doit être installée sur le serveur de chaque hôpital, et accessible uniquement par le personnel de l'hôpital. Pour rendre les dossiers disponibles instantanément et accessible tout le temps, nous avons proposé de centraliser le système dans le Cloud. Ainsi il bénéficie de tous les avantages d'une application Cloud (accès ubiquitaire et facile, haute disponibilité, maintenance et mise à jour facile, …).

- les dossiers médicaux ne sont accessibles qu'en présence du patient ; Ceci pose problème en situation d'urgence. Nous avons proposé une solution pour gérer cette situation en permettant au professionnel de santé d'accéder au dossier médical d'un patient même quand celui-ci n'est pas présent.

- l'intégrité des données n'est pas vérifiée, étant donné l'importance de ce type de données pour les décisions médicales ; Le médecin doit être sûr de l'exactitude des données avant de les utiliser. Nous proposons l'utilisation des signatures numériques pour vérifier l'intégrité des données médicales.

Notre système respecte toutes les exigences fondamentales de sécurité (cf. Tableau 3.2) :

La confidentialité est assurée par l'anonymat des patients ; L'accès aux dossiers médicaux est sûr. Il est basé sur les cartes de santé (quand le patient est présent). Celles-ci stockent les informations nécessaires pour l'accès aux données (les pseudonymes cryptés et leurs clés de décryptages). Pour améliorer la sécurité en situations d'urgence, la gestion des clés se fait à partir de l'hôpital.

Notre solution résout le problème des menaces internes car les informations non cryptées (informations médicales) sont anonymes et les informations sensibles (les pseudonymes) sont cryptées; et les clés de décryptage ne sont pas stockées sur le cloud.

L'intégrité des données est assurée par les signatures numériques. Une empreinte est créée pour chaque dossier médical, et vérifiée avant chaque utilisation.

Notre solution est déployée sur le cloud ce qui offre une haute disponibilité. L'exigence de disponibilité est donc respectée. Cependant, étant donné que les dossiers sont accessibles via internet, leur disponibilité dépend de la connexion internet. Nous avons résolu ce problème en stockant également les dossiers médicaux (non anonymes) sur le serveur de l'hôpital ainsi les dossiers restent accessibles mêmes s'il n'y a pas de connexion internet.

Respect des exigences de sécurité			Méthode utilisée pour assurer la Confidentialité	Méthode utilisée pour assurer l'intégrité	Méthode utilisée pour gérer les situations d'urgence
C	I	D			
Oui	Oui	Oui	Anonymat des patients, Cryptage des informations nécessaires pour l'accès aux données et le stockage des clés de décryptage sur le serveur de l'hôpital, Authentification par cartes	Signatures numériques	Récupération des informations nécessaires pour l'accès aux données du Cloud

Tableau 3.2: Respects des exigences fondamentales de sécurité.

3.6 Conclusion

Nous avons proposé une architecture pour sécuriser les données médicales sur le Cloud. Nous nous sommes basés sur le système P^3HR proposé par [66] que nous avons amélioré. Nous avons décrit l'architecture de notre système ainsi que son fonctionnement général. D'après l'analyse de nous avons faite, notre architecture respecte toutes les exigences fondamentales de sécurité à savoir : la confidentialité, l'intégrité, et la disponibilité.

Par rapport au système proposé par les auteurs, notre système est plus facile à implémenter, à utiliser et à maintenir car c'est une application web. Il facilite le partage et l'échange d'informations entre les professionnels de santé grâce à la centralisation des dossiers médicaux sur le Cloud. Avec notre architecture, les dossiers médicaux deviennent accessibles tout le temps et de n'importe où rapidement. De plus, notre système est performant car il n'utilise pas le cryptage pour assurer la confidentialité des données médicales.

CHAPITRE 4
EXPERIEMTATION ET VALIDATION

4.1 Introduction

Nous avons proposé dans le chapitre précédent une architecture pour sécuriser les données médicales sur le Cloud. Notre architecture est composée de deux parties principales : l'infrastructure de l'hôpital et l'infrastructure Cloud. Pour tester notre solution nous avons réalisé deux applications (*EHR-Manager*, *CEHR-Manager*) représentant chacune une partie de l'architecture. Nous allons décrire dans cette partie ces deux applications, quelques détails d'implémentation et les structures de données utilisées.

4.2 Application 1 : Electronic Health Record Manager (EHR-Manager)

EHR-Manager est une application standard s'exécutant sur le serveur de l'hôpital. Elle est responsable de la manipulation de la base de données, la création des pseudonymes, de la base de données anonyme et des empreintes pour vérifier l'intégrité des dossiers médicaux.

4.2.1 Structure de données

Pour stocker les dossiers médicaux des patients nous avons choisi une base de données relationnelle. Notre base de données est composée de 3 tables : la table Patient, la table Médecin et la table Ehr (Electronic Health Record). Le modèle de données de notre base de données est le suivant :

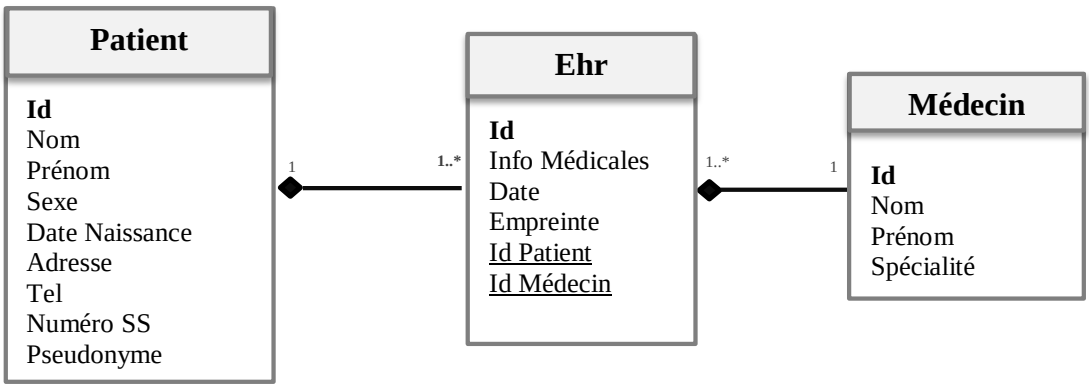

Figure 4.1: Structure de données de la base de données (stockée sur le serveur de l'hôpital).

Pour stocker et interagir avec notre base de données nous avons utilisé le SGBD MySql.

La base de données de l'hôpital est donc constituée des dossiers médicaux des patients contenant leurs informations personnelles et médicales (représentées par une simple description de la pathologie). (cf. Figure 4.2).

Base de données hôpital

ID	NOM	PRENOM	PSEUDONY...	N_SS	S...	DATE_NAI...	ADRESSE	TEL	INFO_MEDICALE	DATE	MEDECIN_N...	MEDECIN_P...	EMPREINTE
1	Ghita	Nora	1106621162	5540070013/58	F	02-07-1955	Zighoud youssef Constantine	0560-12-34-56	Douleurs thoraciques	15-05-2014	Benazouz	Zahra	4a19ad63a6e50...
2	Ghita	Nora	1106621162	5540070013/58	F	02-07-1955	Zighoud youssef Constantine	0560-12-34-56	Fracture de la jambe	23-03-2015	Bouali	Omar	1025ffc629ba12...
3	Ghita	Nora	1106621162	5540070013/58	F	02-07-1955	Zighoud youssef Constantine	0560-12-34-56	Fracture du doigt à la suite à un accident de la route	01-12-2014	Bouali	Omar	a8c05cf7b6e43a...
4	Chahla	Salima	1746520937	8540070013/58	F	09-10-1985	Rue larbi benmhidi Constan...	0555-13-50-38	Fracture du bras suite à un accident de la route	01-12-2014	Bouali	Omar	10d78d9bd1ce5...
5	Fikri	Said	395504976	4840070013/58	M	17-07-1948	Rue fronton Constantine	0551-57-70-65	Fracture du stérnum	01-12-2014	Bouali	Omar	91c4368581263...
6	Fikri	Said	395504976	4840070013/58	M	17-07-1948	Rue fronton Constantine	0551-57-70-65	Violente douleur au ventre	12-11-2013	Zidan	Malik	1aa38f63c98b58...
7	Fikri	Said	395504976	4840070013/58	M	17-07-1948	Rue fronton Constantine	0551-57-70-65	Déséquilibre glycémique	01-01-2012	Zidan	Malik	a9b7e730ae0d9...
8	Hasa	Hichem	1404789609	7640070013/58	M	16-02-1976	Rue abane ramdam Consta...	0661-46-02-66	Crise d'asthme sévère	06-03-2011	Bouhbel	Amir	bd88daaff9a420...
9	Maysan	Malika	1706234378	3040070013/58	F	15-05-1930	Zouaghi constantine	0777-45-67-67	Maladie de parkinson	02-02-2006	Alouane	Warda	9931185331360...
10	Miloud	Lamia	2047158681	7740070013/58	F	24-09-1977	Sidi mabouk constantine	0772-45-67-06	Fracture du col fémoral suite à un accident de la route	01-11-2003	Bouali	Omar	682f72238d75cd...
11	Abbad	Djamel	284720968	8840070013/58	M	24-09-1988	Rue abane ramdam Consta...	0775-45-67-06	Forte grippe	01-06-2015	Zidan	Malik	d33749fdb96904...
12	Rahmou...	Houria	878798554	3040070013/52	F	22-03-1930	Sidi mabrouk constantine	0554-45-67-67	Trouble du rythme cardiaque	01-05-1990	Benazouz	Zahra	6922bc461a447...
14	Chahla	Salima	1746520937	8540070013/58	F	09-10-1985	Rue larbi benmhidi Constan...	0555-13-50-38	Appendicite aigue	01-01-1988	Himore	Ali	4511a537a451e...
18	Chahla	Salima	1746520937	8540070013/58	F	09-10-1985	Rue larbi benmhidi Constan...	0555-13-50-38	Anémie sève	01-04-2015	Zidan	Malik	86b95b4204a0b...
19	Miloud	Lamia	2047158681	7740070013/58	F	24-09-1977	Sidi mabouk constantine	0772-45-67-06	Vomissement, maux de ventre	10-10-2015	Zidan	Malik	5d8c54d3bda80...
20	Miloud	Lamia	2047158681	7740070013/58	F	24-09-1977	Sidi mabouk constantine	0772-45-67-06	Fracture du bras	20-04-2015	Bouali	Omar	49ec6992fb0740...
21	Abbad	Djamel	284720968	8840070013/58	M	24-09-1988	Rue abane ramdam Consta...	0775-45-67-06	Migraines aigues	10-10-2015	Zidan	Malik	6f2f7ccb0a690d...
22	Abid	Hanane	479784036	8640070013/55	F	11-01-1986	Massinissa el khroub	0555-11-11-12	Méningite cérébro-spinal	12-10-2014	Alouane	Warda	1339e772811c1...
23	Abid	Hanane	479784036	8640070013/55	F	11-01-1986	Massinissa el khroub	0555-11-11-12	Infection urinaire	10-20-2013	Dali	Salim	caf63052a0b921...
25	Sekou	Houda	581936148	8840070013/55	F	28-08-1988	Ain smara constantine	0554-30-20-33	Bronchiolite	10-10-2010	Bouhbel	Amir	52d0950beae56...
32	Tamara	Assia	321972332	8640070013/52	F	14-03-1986	St jean constantine	0560-60-60-60	Le patient a subit une appendicectomie	10-10-2014	Himore	Ali	21639bcc427d6...
80	Yougorta	Amine	1547780361	8240070013/58	M	12-09-1982	El hama constantine	0661-06-66-67	Maladie alzheimer	20-02-2011	Zaidi	Leila	7b9c1edab434f6...

Figure 4.2: Base de données de l'hôpital.

4.4.2 Implémentation et description de l'application EHR-Manager

EHR-Manager a été réalisée en Java Standard Edition en utilisant l'IDE Eclipse Java SE. L'interface graphique de l'application est représentée dans la figure 4.3.

Figure 4.3: Interface graphique EHR-Manager.

L'application *EHR-Manager* est constituée de trois modules principaux : le module d'interaction avec la base de données, le module de contrôle de l'intégrité et le module de contrôle de la confidentialité.

4.4.2.1 Module d'interaction avec la base de données

Ce module est responsable de l'interrogation de la base de données des dossiers médicaux de l'hôpital. Il permet la consultation, l'ajout, la modification et la suppression des dossiers médicaux.

- Consultation d'un dossier médical

Ce composant permet la consultation d'un dossier médical et la vérification de sa validité. Pour consulter le dossier médical d'un patient, l'utilisateur doit d'abord l'identifier en introduisant son numéro de sécurité sociale ou son pseudonyme. Le dossier médical correspondant au patient est affiché, et l'utilisateur peut vérifier sa validité. (cf. Figure 4.4).

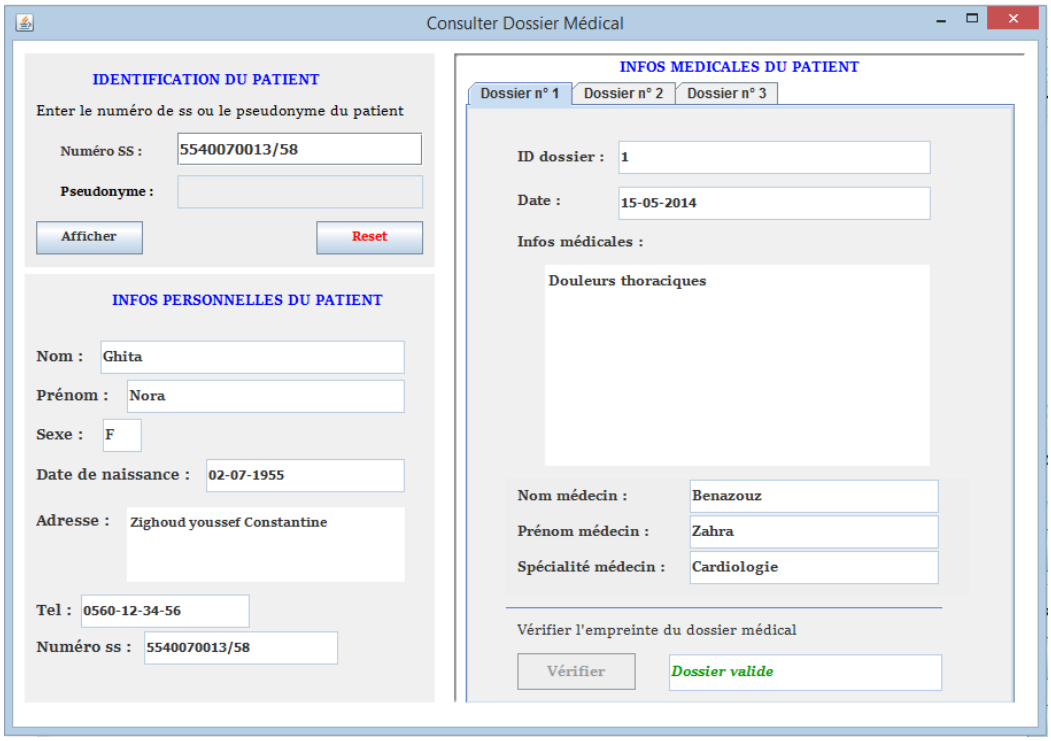

Figure 4.4: Consultation et vérification de la validité d'un dossier médical.

➢ Modification d'un dossier médical

Ce composant permet la modification d'un dossier médical. L'utilisateur (généralement le médecin) peut modifier directement le dossier médical d'un patient après avoir identifié ce dernier avec son numéro de sécurité sociale ou son pseudonyme. Après la modification du dossier, l'utilisateur doit lui créer une nouvelle empreinte. (cf. Figure 4.5)

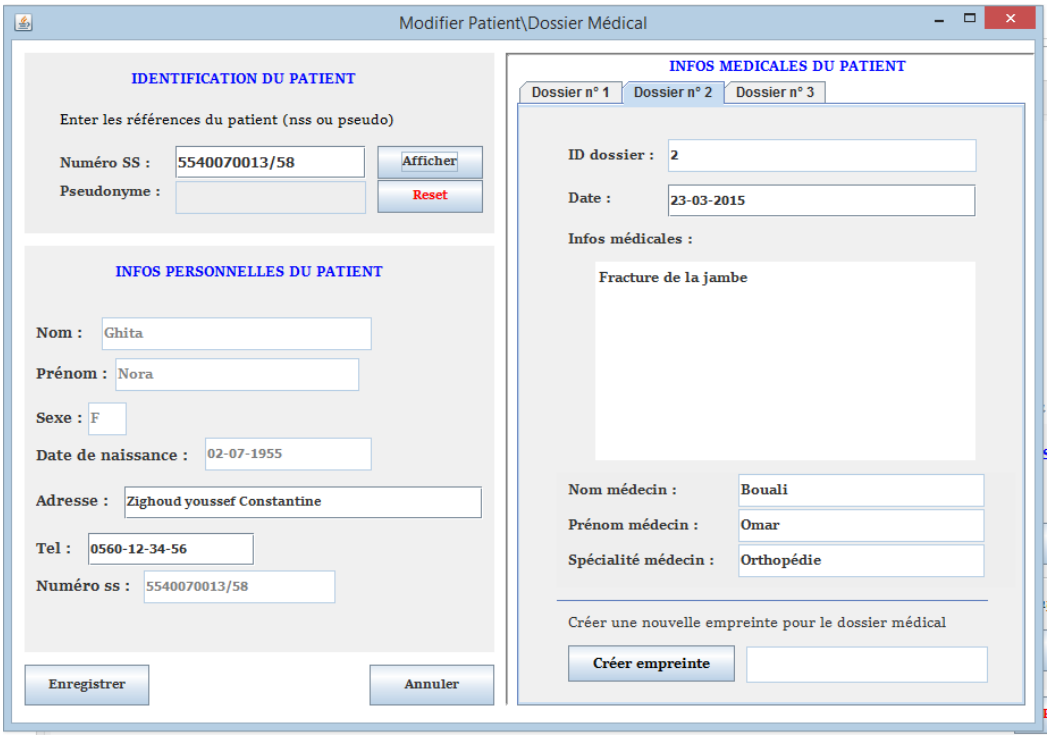

Figure 4.5: Modification et vérification de la validité d'un dossier médical.

➢ Ajout d'un nouveau patient/dossier médical

Ce composant permet soit l'ajout d'un dossier médical pour un patient existant dans la base de données, soit la création d'un nouveau patient (cf. Figure 4.6).

Les figures 4.7 et 4.8 représentent respectivement l'ajout d'un dossier médical à un patient existant et la création d'un nouveau patient.

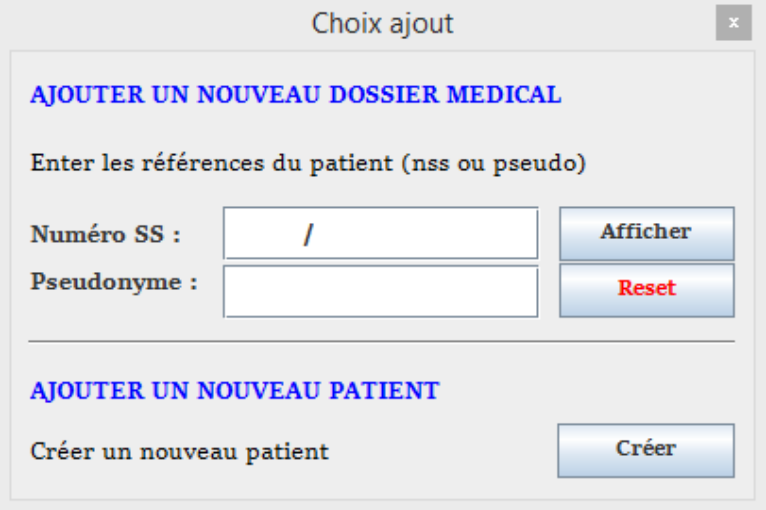

Figure 4.6: Ajout d'un nouveau patient/dossier médical.

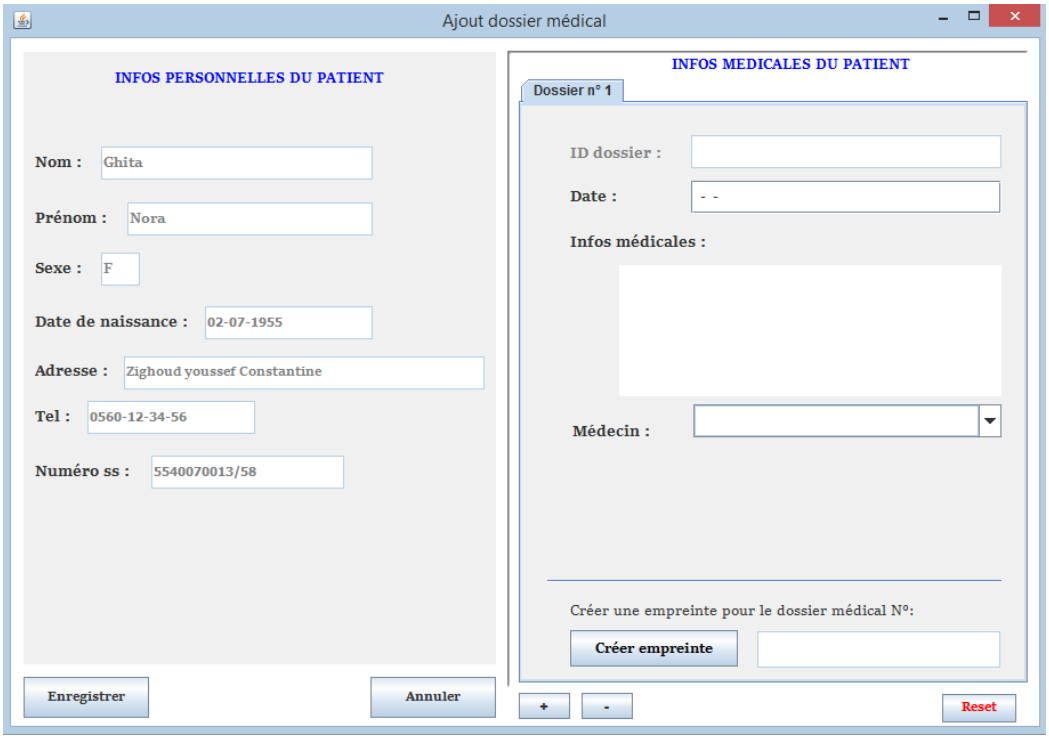

Figure 4.7: Ajout d'un dossier médical à un patient existant.

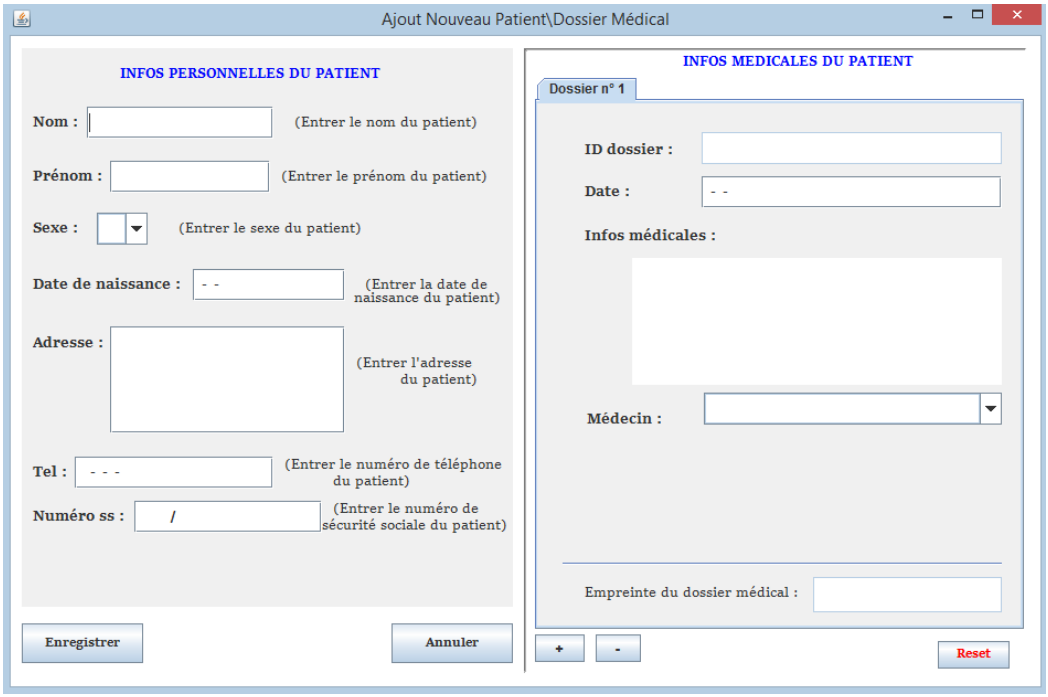

Figure 4.8: Ajout d'un nouveau patient.

- Suppression d'un dossier médical

Ce composant permet soit, la suppression du dossier médical d'un patient en identifiant le dossier à supprimer avec son Id, soit la suppression d'un patient (suppression de tous ses dossiers médicaux) en identifiant le patient avec son numéro de sécurité sociale ou son pseudonyme. (cf. Figure 4.9)

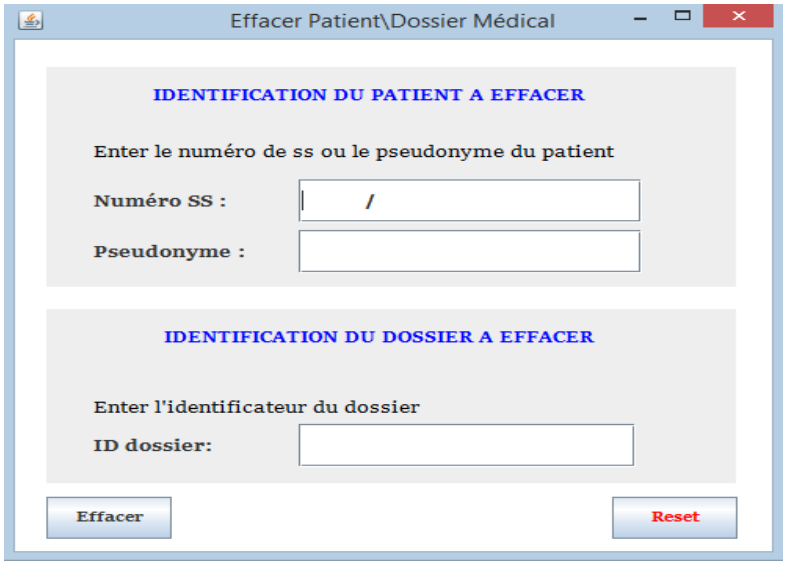

Figure 4.9: Suppression d'un patient/dossier medical.

4.4.2.2 Module de contrôle de la confidentialité

Ce module est responsable de la création et de l'enregistrement des pseudonymes dans la base de données. Le pseudonyme d'un patient est créé à partir de son numéro de sécurité sociale (nss) comme suit : le nss est concaténé avec un nombre aléatoire compris entre 0 et 1 et crypté avec la fonction RSA. Le résultat du cryptage est ensuite haché avec la méthode hash code (pour avoir des pseudonyme entiers et de taille réduite). Les classes responsables de la création des pseudonymes et du cryptage sont les suivantes :

```
public class Pseudo {

    static int pseudonyme=0;

    public static int generer(String nss) {
        MyRSA rsa=new MyRSA();
        rsa.genererCle();
        pseudonyme=new BigInteger(rsa.crypt(nss+Math.random())).hashCode();
        if (pseudonyme<0)
            pseudonyme=pseudonyme*(-1);
        return pseudonyme;
    }
}
```

Algorithme 4.1 : Création des pseudonymes.

```
public class MyRSA {
    public final static int taille_cle = 1024;
    private RSAPublicKey cle_pub;
    private RSAPrivateKey cle_priv;
    public MyRSA() {
            }
    public void genererCle() {
        try {
            KeyPairGenerator cle_generateur = KeyPairGenerator.getInstance("RSA");
            cle_generateur.initialize(taille_cle, new SecureRandom());
            KeyPair cles = cle_generateur.generateKeyPair();
            cle_pub = (RSAPublicKey)cles.getPublic();
            cle_priv = (RSAPrivateKey)cles.getPrivate();
            }
        catch (Exception e) {
            System.out.println("erreur avec la génération des clés");
            }
        }
    private BigInteger crypt(BigInteger msg) {
        return msg.modPow(cle_pub.getPublicExponent(),cle_pub.getModulus());
        }
    private BigInteger decrypt(BigInteger msg) {
        return msg.modPow(cle_priv.getPrivateExponent(), cle_priv.getModulus());
      }
    public byte[] crypt(byte[] text) {
        return crypt(new BigInteger(text)).toByteArray();
        }
    public byte[] crypt(String plaintext) {
        return crypt(plaintext.getBytes());
        }
    public byte[] decryptInBytes(byte[] ciphertext) {
        return decrypt(new BigInteger(ciphertext)).toByteArray();
        }
    public String decryptInString(byte[] ciphertext) {
        return new String(decryptInBytes(ciphertext));
        }
    }
```

Algorithme 4.2 : La classe RSA.

4.4.2.3 Module de contrôle de l'intégrité

Ce module est responsable de la création et de l'enregistrement des empreintes dans la base de données. Les empreintes sont créées avec la fonction de hachage SHA-256 comme suit :

```
public class Empreinte {

    String empreinte;

    public static String generer(String dossier) {

        String  empreinte=MySHA.hacher(dossier.getBytes());

        return empreinte;

        }
    }
```

Algorithme 4.3 : Création des empreintes.

```
public class MySHA {
    public MySHA() {

    }
    public static String hacher(byte[] msg) {

        MessageDigest md;

        try {
            md = MessageDigest.getInstance("SHA-256");
            md.update(msg, 0, msg.length);
            return new BigInteger(1, md.digest()).toString(16);
        } catch (NoSuchAlgorithmException e) {
            e.printStackTrace();
        }
        return null;

    }

    }
```

Algorithme 4.4 : La classe SHA-256.

Nous avons utilisé dans cette application une authentification basée sur les mots de passe (notre application est accessible avec le mot de passe « magister »). (cf .Figure 4.10).

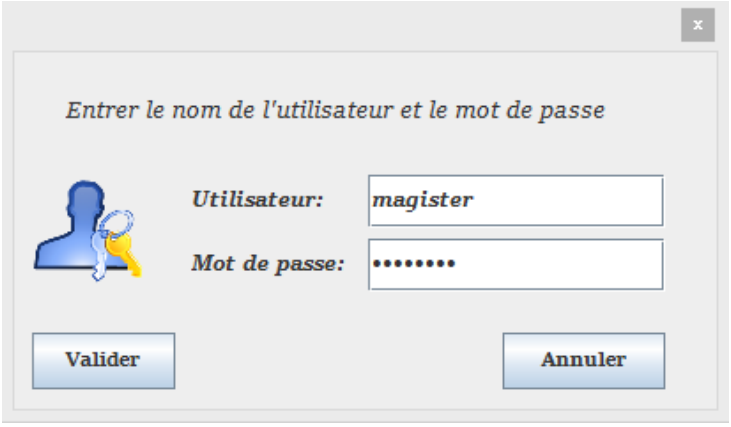

Figure 4.10: Authentification des utilisateurs.

L'objectif final de cette première application est la création de la base de données anonyme contenant les pseudonymes des patients avec leurs informations médicales et l'empreinte de chaque dossier comme le montre la figure suivante :

Base de données anonyme

ID	PSEUDONYME	INFO_MEDICALE	DATE	MEDECIN_N...	MEDECIN_PR...	EMPREINTE
1	1106621162	Douleurs thoraciques	15-05-2014	Benazouz	Zahra	4a19ad63a6e50dabcbdbbd36eff67b89fc3b4331b72464e24685d43e1428017d
2	1106621162	Fracture de la jambe	23-03-2015	Bouali	Omar	1025ffc629ba121d058a6d0b359853d30b21d05dcc92fb81c38acd7b4798036
3	1106621162	Fracture du doigt à la suite à un accident de la route	01-12-2014	Bouali	Omar	a8c05cf7b6e43a148db754e3790f52c79860f5c81352d97b471bce20b0e382c4
4	1746520937	Fracture du bras suite à un accident de la route	01-12-2014	Bouali	Omar	10d78d9bd1ce55dea9fbd10b6f3970e3ee4b5ce1dcf01b11d4ee8f170857f72f
5	395504976	Fracture du stérnum	01-12-2014	Bouali	Omar	91c4368581263d3dd635361f9a14c0b1b709901a10f9af770dd7b7f5c1fe68d7
6	395504976	Violente douleur au ventre	12-11-2013	Zidan	Malik	1aa38f63c98b582927dd78b49684f55571de511aff4bf046b752290444c0d9b7
7	395504976	Déséquilibre glycémique	01-01-2012	Zidan	Malik	a9b7e730ae0d9ee36d6074ddd57793244165990ceb804771aa2cf51179b7195f
8	1404789609	Crise d'asthme sévère	06-03-2011	Bouhbel	Amir	bd88daaff9a420165dd974cdd67ab26a8a7e801df3f709668e64c78caaba2db
9	1706234378	Maladie de parkinson	02-02-2006	Alouane	Warda	993118533313604cf8554f0f7a6e7c2662bd06fcafefcaff257688227c5d4166c
10	2047158681	Fracture du col fémoral suite à un accident de la route	01-11-2003	Bouali	Omar	682f72238d75cd466fe520720a75412ff8bf2dfac7f050ed6413b9cc4dde0075
11	284720968	Forte grippe	01-06-2015	Zidan	Malik	d33749fdb96904ba13da5200eae6320eb61da533a500f9bd50065281f9af0cf4
12	878798554	Trouble du rythme cardiaque	01-05-1990	Benazouz	Zahra	6922bc461a4477df0fbf9a53036c32afcb4bafdd7facf8bff0b5a989b10ecc55
14	1746520937	Appendicite aigue	01-01-1988	Himore	Ali	4511a537a451ee37824e5b8b7a724505de6d9e1bf84989e766837657a71bcc9b
18	1746520937	Anémie sève	01-04-2015	Zidan	Malik	86b95b4204a0b52359ba6d212d864705acebd1a05cfb34c19d983ff86c9e3653
19	2047158681	Vomissement, maux de ventre	10-10-2015	Zidan	Malik	5d8c54d3bda80c2616b292933c802cb5be8a777abfa20dedeadca8999da23d8
20	2047158681	Fracture du bras	20-04-2015	Bouali	Omar	49ec6992fb0740c7d3282cb286c7577c81c67d1ca5f6be6ce410d62f061ea337
21	284720968	Migraines aigues	10-10-2015	Zidan	Malik	6f2f7ccb0a690de7ab91f5495dd78d267db95a96ca6d2386eca05579e09bafe2
22	479784036	Méningite cérébro-spinal	12-10-2014	Alouane	Warda	1339e772811c18e4d7e3cc9a92021a3803865a05eaff1ae263fe2eea4b8fc672
23	479784036	Infection urinaire	10-20-2013	Dali	Salim	caf63052a0b921dc4ced20c0b1d1ecd34d1374f1535838a88e63e2130da0bd76
25	581936148	Bronchiolite	10-10-2010	Bouhbel	Amir	52d0950beae56ed7077a8dabb1d96408e218cb99bf4d0dfce6064a4f1ae9140b
32	321972332	Le patient a subit une appendicectomie	10-10-2014	Himore	Ali	21639bcc427d62c8f535ec656d60cde38f7a30e4610553b31d02ce38718cecc9
80	1547780261	Maladie alzheimer	20-02-2011	Zaidi	Leila	7b0c1edab43ff6d65013775403821254 20fb4f9ad1a1066934b1eb44d7bfb2f2

Figure 4.11: Base de données anonyme.

La prochaine étape consiste à stocker et à accéder à cette base de données (anonyme) dans le cloud, et c'est le rôle de la deuxième application.

Remarque : le pseudonyme du patient est l'élément qui fait le lien avec son dossier médical. Etant donné que notre mécanisme de confidentialité repose sur « non connaissance » de ce lien, celui-ci doit rester secret. Contrairement au Cloud, l'hôpital est considéré comme un environnement sécurisé, et c'est là où sont créés les pseudonymes. C'est pour cette raison que nous avons utilisé les pseudonymes pour chercher les dossiers médicaux des patients. Dans le cloud les pseudonymes ne doivent pas être révélés, et ne peuvent donc pas être utilisés pour chercher les dossiers des patients. Les recherches se feront avec leurs numéros de sécurités sociales qui contrairement aux pseudonymes ne sont pas confidentiels, et peuvent être connus par tous. L'information à protéger dans le Cloud est donc la correspondance entre les pseudonymes et les numéros de sécurité sociale.

4.3 Application 2 : Cloud Electronic Health Record Manager (CEHR-Manager)

CEHR-Manager est une application web s'exécutant sur le Cloud (SaaS), permettant la consultation de la base de données anonyme et la vérification de son intégrité.

4.3.1 Implémentation de l'application CEHR-Manager

CEHR-Manager a été développée en java Entreprise Edition (Java EE) avec Eclipse, et déployée sur la plateforme de Google (Google App Engine). Google App Engine est un service gratuit jusqu'à 1go de stockage ; Il est donc suffisant pour tester notre application. Google fournit un kit de développement App Engine sous la forme d'un plugin Eclipse qui permet de développer et de déployer des applications sur les serveurs de Google. L'outil permet également de tester le fonctionnement d'une application en simulant un serveur Google localement.

Après l'installation du plugin l'icône Google apparait dans Eclipse qui donne accès aux fonctionnalités de Google App Engine. (cf. Figure 4.12)

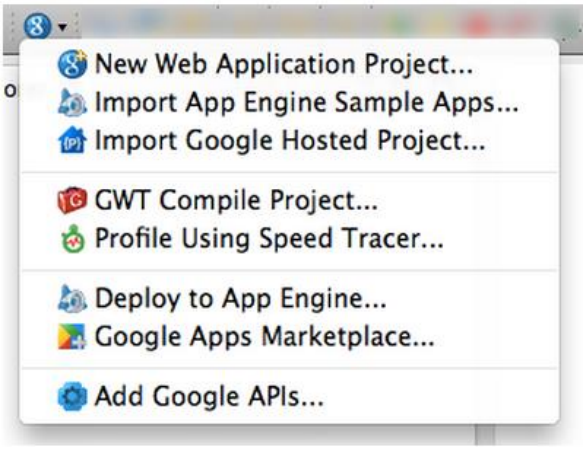

Figure 4.12: Le menu Google dans Eclipse.

Après avoir créé notre application web, nous l'avons déployée sur la plateforme de Google en suivant les étapes suivantes : nous avons, dans un premier temps, réservé un identifiant d'application libre pour notre application sur appengine.google.com, nous avons choisi « cehrconsult ». Nous avons ensuite, dans un second temps, déployé notre application directement depuis Eclipse après avoir indiqué l'identifiant de l'application dans le fichier appengine-web.xml (le fichier de configuration de Google App Engine).

Pour stocker notre base de données anonyme, nous avons besoin d'un serveur de base de données. Google fournit un service appelé Google Cloud SQL qui consiste en une base de données MySQL gérée dans le Cloud. Etant donné que notre base de données anonyme est une base de données relationnelle, nous n'avons donc pas besoins de changements, nous avons juste à utiliser la base de données MySQL de Google qui sera compatible ; Malheureusement ce service est payant, même pour de petites quantités de données.

Nous avons choisi comme alternative à Google Cloud SQL, d'utiliser un autre service de stockage : le Datastrore. Le Datastore est l'outil conseillé par Google. Il est facile d'utilisation, très rapide et peut être utilisé gratuitement (jusqu'à une certaine limite de trafic). Il est basé sur Big Table (le système de stockage utilisé

par Google lui-même pour stocker toutes ses données), et fonctionne en mode High Replication Datastore (HRD), il est donc très résilient aux erreurs.

Le seul « inconvénient » pour nous est que le Datastore n'est pas une base de données relationnelle ; Nous ne pouvons donc pas utiliser notre base de données MySQL directement (comme ça aurait été le cas si nous avions utilisé Google Cloud SQL). Le Datastore se base sur le concept d'entités pour structurer les données. Les entités sont des regroupements d'une ou plusieurs paires clé-valeur. Pour utiliser ce concept, nous avons dû recréer notre base de données anonyme sur Google App Engine.

4.3.2 Nouvelle structure de données

Notre base de données anonyme est composée de deux tables (la table Ehr et la table Pseudo). Etant donné que les identités des patients ont été supprimées, la table Patient n'existe plus. Nous avons également enlevée la table Médecin dans un souci de simplification.

Pour stocker notre base de données anonyme dans le Datastore nous avons besoins de créer une entité par table comme suit :

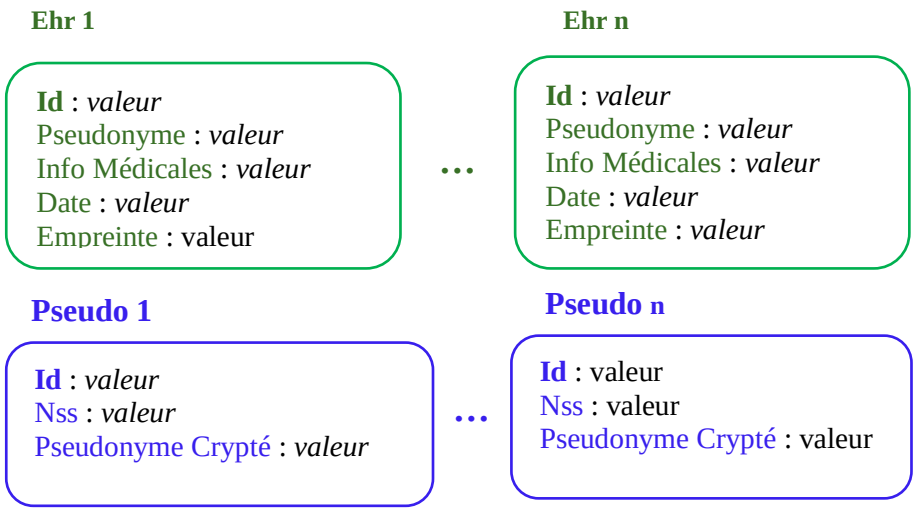

Figure 4.13: Structure de données de la base de données anonyme stockée sur dans le Datastore Google.

Etant donné que le lien entre le patient et son dossier médical a été caché, il faut trouver un moyen pour faire correspondre un patient à son dossier, sans révéler ce lien. Pour cela nous avons ajouté l'entité Pseudo dans le Datastore qui contient le numéro de sécurité sociale du patient et le pseudonyme correspondant crypté. Le cryptage des pseudonymes assure la confidentialité du lien entre un patient et son dossier médical, car personne à part lui (et les personnes autorisées) ne détient la clé de décryptage, et ne peut donc décrypter le pseudonyme.

Dans notre implémentation, nous avons utilisé des pseudonymes non cryptés, dans la mesure où nous n'avions pas les cartes de santé qui stockent les clés de décryptage pour tester notre application. La figure 4.14 représente l'entité Pseudo de la base de données anonyme dans le Datastrore.

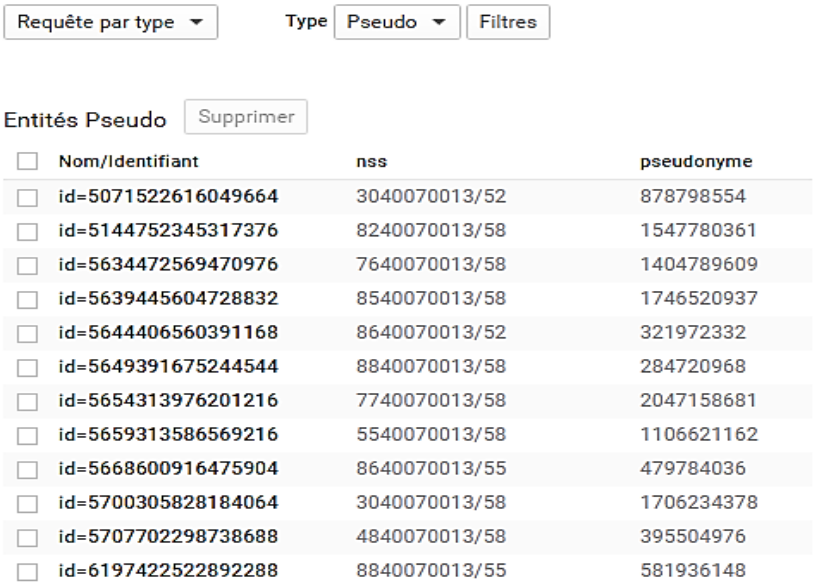

Figure 4.14: Stockage de l'entité Pseudo dans le Datastore.

4.3.3 Description de l'application CEHR-Manager

L'interface graphique de l'application *CEHR-Manager* est représentée dans la figure 4.15. *CEHR-Manager* est une application SaaS s'exécutant sur Cloud et accessible via internet à l'adresse suivante: *http://cehrconsult.appspot.com/cehrmanager.*

Figure 4.15: Page d'accueil de CEHR-Manager.

CEHR-Manager permet la consultation des dossiers médicaux anonymes sur le Cloud. La recherche se fait avec le numéro de sécurité sociale du patient, qui sera utilisé pour chercher le pseudonyme correspondant. Etant donné que le pseudonyme doit rester secret il ne doit pas être affiché avec le dossier médical. (cf. Figure 4.16)

Figure 4.16: Recherche et affichage d'un dossier médical.

L'utilisateur peut vérifier la validité du dossier médial affiché en cliquant sur « Vérifier » (cf. Figure 4.17)

Figure 4.17: Vérification de la validité d'un dossier médical.

Nous avons utilisé dans cette application une authentification basée sur les mots de passe (notre application est accessible avec le mot de passe « magister »). (cf. Figure 4.18)

Figure 4.18: Authentification de CEHR-Manager.

Google fournie une interface « APIs Console Google » qui permet de contrôler l'application en se connectant à internet (elle permet entre autre d'accéder à l'espace de stockage : Datastore). L'API est accessible à l'adresse https://console.developers.google.com.

4.4 Conclusion

Le travail rapporté dans ce chapitre a consisté essentiellement à décrire l'implémentation de notre architecture ainsi que les outils utilisés. Le résultat de ce chapitre est que nous avons sécurisé nos données médicales sur le Cloud : nos données médicales sont confidentielles, disponibles sans interruption et leurs

validité peut être vérifiée. Nous pensons que l'objectif final de ce travail a été atteint avec ce dernier chapitre à savoir la proposition d'une solution pour protéger les données médicales sur le Cloud.

CONCLUSION

L'objectif de ce travail est de proposer une solution pour sécuriser les données médicales sur le Cloud Computing. Notre contribution consiste en une architecture de sécurisation qui se résume principalement en une application Cloud hébergée sur la plateforme de Google, et qui présente les caractéristiques suivantes:

Premièrement, les dossiers médicaux des patients sont centralisés dans le Cloud. Ils sont donc disponibles tout le temps, et accessibles par n'importe quel hôpital, ce qui facilite l'échange et le partage d'informations entre les différentes institutions médicales.

Deuxièmement, notre architecture est une application Cloud, elle est donc facile à réaliser et à maintenir.

Enfin, l'utilisation du Cloud décharge l'institution médicale de la gestion technique, et peut se consacrer à son objectif principal : la santé des patients.

Nous avons démontré dans les chapitres précédents que notre architecture répond à toutes les exigences fondamentales de sécurité requises pour les données médicales, à savoir : la confidentialité, l'intégrité et la disponibilité : La confidentialité est assurée par l'anonymat des patients et le cryptage des pseudonymes stockés sur le cloud. L'intégrité des données est assurée par les signatures numériques, et la disponibilité par le stockage de base de données dans le Cloud et sur le serveur de l'hôpital. Nous avons également proposé une solution pour gérer les situations d'urgences.

Notre solution est performante car elle se base sur l'anonymat et non sur le cryptage pour assurer la confidentialité des données. Nous utilisons le cryptage uniquement pour crypter les pseudonymes (quantité des données cryptée est faible).

Toutefois notre solution reste théorique dans l'ensemble car elle n'a pas été expérimentée sur des données réelles. De plus Notre solution n'est pas achevée. Nous envisageons de lui apporter les améliorations suivantes :

- utilisation d'un contrôle d'accès basé sur les cartes de santé (en utilisant les cartes CHIFA par exemple).
- expérimentation de notre solution sur une base de données concrète.
- l'enregistrement des dossiers médicaux anonymes a été fait manuellement. Il restera à ajouter un module d'enregistrement dans le Cloud.
- l'implémentation des signatures des dossiers médicaux avec les cartes des professionnels de santé.

REFERENCES

1. S.Ramgovind, M.M.Eloff, E.Smith. *"The Management of Security in Cloud Computing"*. IEEE Information Security for South Africa (ISSA). Août 2010.

2. Philogene A. Boampong, Luay A. Wahsheh *"Different Facets of Security in the Cloud".* ACM Proceedings of the 15th Communications and Networking Simulation Symposium CNS '12. Mars 2012.

3. Chunming Rong, Son T. Nguyen, Martin Gilje Jaatun *"Beyond lightning: A survey on security challenges in cloud computing".* ScienceDirect Computers and Electrical Engineering. Mai 2012.

4. H. Yu, N. Powell, D. Stembridge, X. Yuan. *"Cloud Computing and Security Challenges".* ACM-SE '12 Proceedings of the 50th Annual Southeast Regional Conference. Mars 2012.

5. Madhan Kumar Srinivasan, K Sarukesi, Paul Rodrigues, Sai Manoj M, Revathy P. *"State-of-the-art Cloud Computing Security Taxonomies-A classification of security challenges in the present cloud computing environment".* International Conference on Advances in Computing, Communication and Informatocs. Août 2012.

6. Aderemi A. Atayero, Oluwaseyi Feyisetan. *"Security Issues in Cloud Computing: The Potentials of Homomorphic Encryption".* Journal of Emerging Trends in Computing and Information Science. Octobre 2011.

7. Pascal Sauliere. *"Cloud Computing et sécurit*é". Cycle de conférence sur le Cloud Computing et Virtualisation. Clubs de la sécurité de l'information régionaux (Clusif). Avril 2010.

8. CloudComputing.fr.

9. Y.Jadeja, K.Madi. *"Cloud Computing- Concepts, Architecture ans Challenges".* IEEE International Conference on Computing, Electronics and Electrical Technologies (ICCEET). Mars 2012.

10. Qi Zhang, Lu Cheng, Raouf Boutaba. *“Cloud Computing : State-of-art and Research Challenges”.* International Journal of Computer Applications (IJCA). Avril 2010.

11. Gowthan Gajala. *“Cloud Computing : A State of art of the Cloud”*. International Journal of Computer Trends and Technology (IJCTT). Janvier 2013.

12. G.Vijay Bhaskar, N.Satheesh Kumar, N.Karthik. *“Research Analysis of Cloud Computing”.* International Journal of Computer Science and Mobile Computing (IJCSMC). Septembre 2013.

13. François Santy. *“La virtualisation”*. Projet de recherche et communication scientifique. Université Libre de Bruxelles. 2009/2010.

14. aws.amazon.com.

15. Windowsazure.com.

16. Google.com.

17. Gridcomputing.com.

18. Sneha Prabha Chandran and Mridula Angepat. *“Cloud Computing: Analysing the risks involved in cloud computing environments”.* Innovation Design and Engineering (IDT) Malardalen University Sweden.Octobre 2010.

19. Keiko Hashizume, David G Rosado, Eduardo Fernández-Medina, et Eduardo B Fernandez. *“An analysis of security issues for cloud computing”.* Journal of Internet Services and Applications (JISA). A SpringerOpen journal. Février 2013.

20. Hyangjin Lee, Jeeyeon Kim, Youngsook Lee, et Dongho Won. *”Security Issues and Threats According to the Attribute”.* Springer Communications in Computer and Information Science. Décembre 2012.

21. Chirag Modi, Dhiren Patel, Bhavesh Borisaniya, Avi Patel, Muttukrishnan Rajarajan. *”A survey on security issues and solutions at different layers of Cloud computing”.* Springer The Journal of Supercomputing. Octobre 2012

22. G.Kulkarni et J.Gambhir, T.Patil, A.Dongare. *”A Security Aspects in Cloud Computing”.* IEEE 3rd International Conference on Software Engineering and Service Science (ICSESS). Juillet 2012.

23. Deyan Chen, Hong Zhao. *“Data Security and Privacy Protection Issues in Cloud Computing”.* IEEE International Conference on Computer Science and Electronics Engineering (ICCSEE). Mars 2012.

24. Abdullah Abuhussein, Harkeerat Bedi, Sajjan Shiva. *“Evaluating Security and Privacy in Cloud Computing Services: A Stakeholder's Perspective”*. IEEE The 7th International Conference for Internet Technology and Secured Transactions (ICITST). Décembre 2012.

25. Eman M.Mohamed, Hatem S. Abdelkader. *” Enhanced Data Security Model for Cloud Computing”.* IEEE The 8th International Conference on INFOrmatics and Systems (INFOS) - Cloud and Mobile Computing Track. Mai 2012.

26. Minqi Zhou, Rong Zhang, Wei Xie, Weining Qian, Aoying Zhou. *” Security and Privacy in Cloud Computing: A Survey”.* IEEE 6th International Conference on Semantics, Knowledge and Grids. Novembre 2010.

27. Sandeep K.Sood. *“A combined approach to ensure data security in cloud computing”.* ScienceDirect Journal of Network and Computer Applications. Novembre 2012.

28. H.Tianfield. *“Cloud Computing Architecture”.* IEEE International Conference on Systems, Man, and Cybernetics (SMC). Octobre 2011.

29. Madhan Kumar Srinivasan, K Sarukesi, Paul Rodrigues, Sai Manoj M, Revathy P. *“State-of-the-art Cloud Computing Security Taxonomies-A classification of security challenges in the present cloud computing environment”.* ACM International Conference on Advances in Computing, Communication and Informatics ICACCI. Août 2012.

30. V.Krishna Reddy, Dr. L.S.S.Reddy. *“Security Architecture of Cloud Computing”.* International Journal of Engineering Science and Technology (IJEST). Septembre 2011.

31. Shilpashree Srinivasamurthy, David Q. Liu *“Survey on Cloud Computing Security”.* IEEE 2nd International Conference on Cloud Computing Technology and Science. 2010.

32. R.Chow, P. Golle, M.Jakobsson, R. Masuoka, J. Molina, E. Shi, J.Staddon. *“Controlling Data in the Cloud: Outsourcing computation without Outsourcing Control”.* ACM Workshop on Cloud Computing Security (CCSW). Novembre 2009.

33.Kaushik Raghupathi. *“5 key events in the history of Cloud Computing”.* Cloud.dzone.com. 2011.

34. salesforce.com.

35. Techopedie.com.

36. Serge RICHARD. *“A propos de la sécurité des environnements virtuels”.* Clubs de la sécurité de l'information régionaux (Clusir).2013.

37. Hyokyung Chang et Euiin Choi. *”Challenges and Security in Cloud Computing”.* Springer International Conference, Future Generation Information Technology Conference, FGIT. Novembre 2010.

38. Ainul Azila Che Fauzi, A. Noraziah, Tutut Herawan, et Noriyani Mohd. Zin. *“On Cloud Computing Security Issues”.* Springer 4th Asian Conference Intelligent Information and Database Systems ACIIDS. Mars 2012.

39. Du meng. *“Data security in cloud computing”.* IEEE The 8th International Conference on Computer Science & Education (ICCSE). Avril 2013.

40. Mark Townsend. *“Managing a Security Program in a Cloud Computing Environment”.* ACM Information Security Curriculum Development Conference InfoSecCD '09 . Septembre 2009.

41. Yubo Tan , Xinlei Wang. *“Research of Cloud Computing Data Security Technology”.* IEEE 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet).Avril 2012.

42. Afnan Ullah Khan, Manuel Oriol, Mariam Kiran, Ming Jiang, Karim Djemame. *“Security Risks and their Management in Cloud Computing”.* IEEE 4th International Conference on Cloud Computing Technology and Science. Décembre 2012.

43. Zhongbin Tang, Xiaoling Wang, Li Jia, Xin Zhang,Wenhui Man. *“Study on Data Security of Cloud Computing”.* IEEE Spring Congress on Engineering and Technology (S-CET). May 2012.

44. S. Subashini n, V.Kavitha. *“A survey on security issues in service delivery models of cloud computing”.* ScienceDirect Journal of Network and Computer Applications. Janvier 2011.

45. Allan A. Friedman, Darrell M. West *“Privacy and Security in Cloud Computing”.* Springer Issues in TECHNOLOGY Innovation. Octobre 2010.

46. Aderemi A. Atayero, Oluwaseyi Feyisetan. *“Security Issues in Cloud Computing: The Potentials of Homomorphic Encryption”.* Journal of Emerging Trends in Computing and Information Science. Octobre 2011.

47. H. Yu, N. Powell, D. Stembridge, X. Yuan. *“Cloud Computing and Security Challenges”.* ACM-SE '12 Proceedings of the 50th Annual Southeast Regional Conference. 2012.

48. Chimere Barron, Huiming Yu et Justin Zhan. *“Cloud Computing Security Case Studies and Research”.* Proceedings of the World Congress on Engineering. Juilet 2013.

49. Amit Sangroya, Saurabh Kumar, Jaideep Dhok, and Vasudeva Varma. *“Towards Analyzing Data Security Risks in Cloud Computing Environments”.* Springer. Mars 2010.

50. Syndicat professionnel de l’écosystème numérique français (Syntec numérique). *“SECURITE DU CLOUD COMPUTING/ Analyse des risques, réponses et bonnes pratiques”.* 2010.

51. Chunming Rong, Son T. Nguyen, Martin Gilje Jaatun *“Beyond lightning: A survey on security challenges in cloud computing”*. SienceDirect Computers and Electrical Engineering. Janvier 2013.

52. Somayeh Sobati moghadam. *“A survey of virtualization security”.* International Journal of Security and Engineering Research. Septembre 2013.

53. F. A. Alvi, B.S Choudary,N. Jaferry , E.Pathan. *“A review on cloud computing security issues & challenges”.* 1st International Conference on Mobility for Life: Technology, Telecommunication, and Problem Based Learning. 2011

54. Siani Pearson. *“Taking Account of Privacy when Designing Cloud Computing Services”.* HP Laboratories. Mars 2009.

55. Kresimir Popovic et Zeljko hocenski. *“Cloud Computing Security issues and challenges”.* Conference publication. Juillet 2010.

56. S. O Kuyoro, F Ibikunle, O Awodele. *“Cloud Computing Security Issues and Challenges”.* International Journal of Computer Networks. Décembre 2011.

57. Jenni Susan Reuben. *”A Survey on Virtual Machine Security”.* Telecommunications software and multimedia laboratory (TML).Seminar on Network Security. Octobre 2007.

58. *“Common Virtualization Vulnerabilities and How to Mitigate Risks”.* Penetration Testing Lab. wordpress.com.

59. *“A Deep Dive Into Hyperjacking”.* SECURITYWEEK NETWORK. securityweek.com.

60. Sara Qaisar et Kausar Fiaz Khawaja. *“CLOUD COMPUTING: NETWORK/SECURITY THREATS AND COUNTERMEASURES”.* Interdiciplinary Journal Of Contemporary Research In Buisness (IJCRB). Janvier 2012.

61. Ajey Singh, Maneesh Shrivastava. *“Overview of Attacks on Cloud Computing”.* International Journal of Engineering and Innovative Technology (IJEIT). Avril 2012.

62. Pao-Ching Chen, Chih-Pin Freg, Ting-Wei Hou, W.-G. Teng. *“Implementing RAID-3 on Cloud Storage for EMR System”.* IEEE International Computer Symposium (ICS). Décembre 2010.

63. Aderonke Justina-Ikuomola, Oluremi O-Arowolo. *“Securing Patient Privacy in E-Health Cloud Using Homomorphic”.* International Journal of Computer Networks and Communications Securit (CNCS). Janvier 2014.

64. Sushmita Ruj, Milos Stojmenovic, Amiya Nayak. *“Privacy Preserving Access Control with Authentication for securing data in the Clouds”.* IEEE/ACM 12th International Symposium on Cluster, Cloud and Grid Computing (CCGrid) .Mai 2012.

65. Zhuo-Rong Li, En-Chi Chang, Kuo-Hsuan Huang, Feipei Lai. *“A Secure Electronic Medical Record Sharing Mechanism in the Cloud Computing platform”.* 15th International Symposium on Consumer Electronics IEEE. Juin 2011.

66. Nurul huda, Noboru Sonehara, Shigeki Yamada. *“A privacy Management Architecture For Patient-controlled Personal Health Record System”.* Journal of Engineering Science and Technology. Juin 2009.

67. Hans Löhr, Ahmad-Reza Sadeghi, Marcel Winandy. *“Securing the E-Health Cloud”.* ACM 1st International Health Informatics Symposium. Novembre 2010.

68. Xueli Huang, Xiaojiang Du. *“Efficiently Secure Data Privacy on Hybrid Cloud”.* IEEE International Conference on Communications (ICC). Juin 2013.

69. Naipeng Dong, Hugo Jonker, Jun Pang. *“Challenges in e-Health: from enabling to enforcing privacy”.* Springer 1st International Symposium. Août 2012.

70. Goce Gavrilov, Vladimir Trajkovik. *“Security and Privacy Issues and Requirements for healthcare Cloud Computing”.* Springer Advances in Intelligent Systems and Computing. Septembre 2013.

71. Fei Liu, Eric Rijnboutt, Dimitrios Routsis. *“What challenges have to be faced when using the Cloud for the e-health service”.* IEEE 15th International Conference on e-Health Networking, Applications and services. Octobre 2013.

72. Assad Abbas, Samee U-Khan. *“A Review on the State-of-the-art Privacy-Preserving Approaches in e-health Clouds”.* IEEE Journal of Biomedical and Health Informatics . Juin 2014.

73. Eman AbuKhousa, Nader Mohamed, Jameela Al-Jaroodi. *“E-Health Cloud-Opportunities and Challenge”.* Journal Future Internet. Juillet 2012.

74. Yu-Yi Chen, Jun-Chao Lu, Jinn-Ke Jan. *“A secure EHR system based on hybrid Clouds”.* Springer Journal of Medical Systems, LLC. Fevrier 2012.

75. Penny Duquenoy, Nermeen Magdi Mekawie, Mark Springett. *“Patients, Trust and Ethics in Information Privacy in eHealth”*. Springer Berlin Heidelberg. 2013.

76. Peter Schartner et Martin Schaffer. *”Unique user-generated digital Pseudonyms”.* Springer Berlin Heidelberg. 2005.

77. Cryptage.org.

78. Damien Riquet, Gilles Grimaud et Michaël Hauspie, « Étude de l’impact des attaques distribuées et multi-chemins sur les solutions de sécurité réseaux », 9ème Conférence Internationale Jeunes Chercheurs, Lille, France, Octobre 2012.