A home healthcare system in the cloud – addressing security and privacy challenges

Mina Deng*, Milan Petković†, Marco Nalin‡ and Ilaria Baroni§
*Philips Research Europe, The Netherlands
Email: mina.deng@philips.com, milan.petkovic@philips.com
†‡§Scientific Institute Hospital San Raffaele, Italy
Email: nalin.marco@hsr.it,ilaria.baroni@hsr.it

Abstract—Cloud computing is an emerging technology that is expected to support Internet scale critical applications which could be essential to the healthcare sector. Its scalability, resilience, adaptability, connectivity, cost reduction, and high performance features have high potential to lift the efficiency and quality of healthcare. However, it is also important to understand specific risks related to security and privacy that this technology brings. This paper focuses on a home healthcare system based on cloud computing. It introduces several use cases and draws an architecture based on the cloud. A comprehensive methodology is used to integrate security and privacy engineering process into the software development lifecycle. In particular, security and privacy challenges are identified in the proposed cloud-based home healthcare system. Moreover, a functional infrastructure plan is provided to demonstrate the integration between the proposed application architecture with the cloud infrastructure. Finally, the paper discusses several mitigation techniques putting the focus on patient-centric control and policy enforcement via cryptographic technologies, and consequently on digital rights management and attribute based encryption technologies.

Keywords—home healthcare; security and privacy; cloud computing; architecture.

I. INTRODUCTION

Cloud computing is one of the emerging technologies that has an increasing impact on both private and public sectors. It represents an on-demand service model for delivering resources ranging from storage and data access, via computation to software provisioning. Typical categories of cloud computing include Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). In the IaaS category, virtual machines are provided, such as Amazon EC2. PaaS offers development tools such as Microsoft Azure, while SaaS providing software on-demand usually over the Internet, eliminating the need for installing and maintenance of the software on the client’s computer, such as Google Docs. The most important benefits of cloud computing include scalability, cost reduction, data availability, reliability and resilience.

There is an acceleration of adoption of cloud computing among enterprises. Gartner projects worldwide cloud service revenues to reach 148.8 billion USD by 2014, with large part from the healthcare cloud computing market. National/regional healthcare authorities as well as healthcare service providers have shown great interests, and are already taking first steps towards the deployment of cloud computing. Cloud computing can help healthcare providers focus more on increasing quality of delivered healthcare instead of managing their IT. This is especially important for smaller hospitals, community care and physician practices. Next to that, cloud computing simplifies information sharing among various healthcare institutions involved in the care process, which is of utmost importance in healthcare. Healthcare cloud has a great market potential given the fact that less than 7% of the US hospitals have a functional and integrated electronic medical record solution.

Unfortunately, cloud computing does not come with benefits only. Moving the infrastructure and sensitive patient data from hospitals to the cloud can pose severe security and privacy risks. Data protection legislation set the rules on how healthcare providers must retain and share patient’s healthcare records. For example, transfer of the records to countries outside the EU without patient’s explicit consent is prohibited unless the data is anonymized. Clouds may involve additional insider attacks or become a single point of failure. These questions are to be addressed by a recent FP7 project TClouds (or Trustworthy Clouds) funded by the European Commission. The project aims to develop an advanced cloud infrastructure that can increase the level of security, privacy and resilience while still being cost-efficient and scalable. Our particular goal in this project is to investigate the use of cloud computing for advanced home healthcare services and develop security and privacy mechanisms for specific healthcare requirements related to cloud computing. This paper presents some of the recent results from the TClouds project and addresses the most privacy and security challenging issues for home healthcare applications in use of a public cloud.

The rest of this paper is organized as follows. Section I.E.A sets the scene and proposes a home healthcare application architecture to be developed in the TClouds project. Section I.E.B analyzes security and privacy threats and elicits requirements for cloud-based home healthcare. Section I.E.C discusses the integration plan to a public cloud and a provisional mitigation solution with a focus on facilitating patient-centric protection for electronic health data. Section I.E.A concludes the paper.
II. ARCHITECTURE OF A HOME HEALTHCARE SYSTEM IN THE CLOUD

A. Home healthcare scenario

In this paper, we consider a cloud-based home healthcare system to empower depressed patients over their treatment process. As depicted in Figure 1, the services offered essentially consists of three components, namely drug therapies management for improving compliance with doctor recommendations, sleep and light management, and physical activity management. In addition, the system also offers a number of services dedicated to healthcare professionals and institutional organizations, based on regional or national healthcare systems.

![Figure 1. TCclouds home healthcare application scenario and actors](image)

The stakeholder involved in the home healthcare scenario include: patients and family members, healthcare professionals (such as (GP) and psychiatrists), pharmacists, health and wellness service providers (such as the one that provides activity monitoring services), and public authority (such as auditing or legal authorities that need to access healthcare data under specific conditions). Personal monitoring devices are used to monitor and collect patient’s physical activity data or sleep information. Besides, the cloud service provider offers platform as a service (PaaS) and infrastructure as a service (IaaS) to host home healthcare applications. In certain cases, delivery service operators are also involved to deliver drugs to the patient’s home.

B. TCclouds home healthcare application architecture

The TCclouds home healthcare application architecture is proposed based on the scenarios in Sec. II-A and graphically represented using data flow diagram (DFD) in Figure 2. The DFD contains four asset types, namely data flows (i.e., data communication), data stores (i.e., database files in repositories), processes (i.e., units of functionality or programs) and external entities (i.e., users or external services). The frames marked with dashed-lines in Figure 2 indicate different business application domains, including Health and Wellness Service Provider, Traditional Healthcare Service Providers (such as hospitals and GP offices), Personal Healthcare Record (PHR) Service Provider, and Regional/National Institutional Service Provider.

Processes can be classified as front end, middle-tier, and back-end. A front end is an interface connecting an external entity with middle-tier. Middle-tiers are business-logic processes that offer various services. Back-ends are the processes that manage data stores.

Taking Figure 2, the middle-tier offered by the home healthcare system are:

- **Prescription management application** (12): for regional/regional authority to manage prescriptions;
- **Epidemiology application** (13): for regional/regional authority to perform epidemiology studies;
- **Prescription application** (17): for healthcare professionals to issue and manage prescriptions;
- **Sleep, drug, and physical activities management application** (18): for healthcare professionals to provide drug management, sleep management, and physical activities management services;
- **Personal diaries and assessment questionnaires management application** (19): for healthcare professionals to manage patients personal diaries and self-assessment questionnaires;
- **Institutional information filter** (20): to filter/anonymize information within healthcare institutions;
- **Personal information filter** (22): to filter/anonymize users’ personal information within the PHR domain;
- **PHR management application** (23): for users to manage their personal healthcare records;
- **Health management application** (27): for health service provide to offer health management services.

There are four types of back-end data stores in the TCclouds home healthcare architecture. The data stored in the **physical activity repository** (31) are subscriber’s monitoring data, such as sleep monitoring and physical activities information, and personal coaching advices provided by health and wellness service providers. The data stored in the PHR repository (30) are patient’s personal health records of patients, including prescriptions, physical activities and sleep information, overview of personal medical and medication information, personal information, and personal privacy policy. The data stored in the EHR repository (29) are electronic health records of patients used for clinical purposes. The data stored in the Prescription repository (28) provide an overview of prescriptions and drug purchase history.

C. Example use case scenario

To instantiate services described in Sec. II-A, during the development process, use cases of the proposed system have
been analyzed, in order to refine the application architecture. Involved actors and data flows in the home healthcare system are highlighted, with the use cases such as patient management portal, personal diary and self assessment questionnaire, physical, sleep, and drug intake activities monitoring, drug therapy management, epidemiological studies and auditing.

Consider one example use case: a depressed patient (6) wears a monitoring device (8) to collect own physical and sleep activities information. The activity data are uploaded from the device, via the health service provider user front end (24) for manual uploading, or the personal device interface (26) for automatic uploading. The data then pass through the middle-tier health management application (27) and back-end management for activity data repository (31), and are stored in the physical activity data repository (35) hosted in the cloud. The stored data may be shared with the PHR service provider on demand.

In order to understand and identify the underlying security and privacy challenges, two strategies have been followed. We name the first one the business-logic driven strategy and the latter the architecture driven strategy. In particular, the requirement elicitation with the business-logic driven strategy relies on brainstorming based on the system functionalities (Sec. II-A, II-C), which is discussed in Sec. III-A. The architecture-driven strategy starts from system assets (i.e. the objects that need to be protected in the system) and analyzes the potential security and privacy threats at these assets, which is discussed in Sec. III-B.

A. Business-logic driven requirements elicitation

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction, as defined by NIST [3]. The cloud has five essential characteristics [3], namely on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. In particular, cloud computing is expected to offer a number of benefits, including the provision of automated self-managed services supporting clouds’ virtual resources availability, multi-tenancy, scalability, resilience, reliability, adaptability, security and privacy, and cost reduction. Apart from these generic requirements of cloud computing, there are a number of specific security and privacy requirements regarding home
healthcare applications in the cloud.

- Semi-trusted cloud service providers: It is assumed that the cloud service providers are semi-trustworthy (honest but curious). This implies that cloud service providers honestly performs legitimate protocols but passively observes traffic in the cloud.

- Data-centric protection: Medical data are stored in highly distributed data stores in the cloud and managed by heterogeneous systems with complex and maybe legacy architectures, some of which may not have a trustworthy data management system. Hence in this setting, data can be potentially accessed by a variable set of parties from different domains with different rights. There is a large uncertainty in who will eventually need to access a data object. It is thus implausible to implement central management. Instead, data protection is decentralized. Moreover, not only data distribution but also data usage should be controlled.

- Emergency access and availability: Timely availability of medical data needs to be guaranteed, especially under emergency cases. This in turn requires the availability of decryption keys if data are encrypted at data stores.

- Efficiency: Data management mechanisms must be sufficiently efficient to suffice the care process and business needs.

- Privacy protection for patients: Due to the sensitive nature, patients’ personal data stored in the cloud need to be protected. In Europe, it is advised that data disclosures are controlled following the data minimization principle \([4]\), namely, the disclosure and retention of personal data should be limited to what is directly relevant and necessary to accomplish a specified purpose. This implies that under certain conditions, medical data needs to be filtered or anonymized according to patient’s privacy policy.

- Patient-centric transparency: Patient should be aware of their privacy rights, and able to specify and delegate the access control policy of their data. Apart from this, accountability should be ensured. Data access and usage need to be logged. In many cases, the context allowing data access cannot be determined automatically, but only verified by a human after the incident. In this regard, auditing should be accompanied with some automated verification procedures.

- Data confidentiality: Unlike multimedia or entertainment data, even partial leakage of patient’s medical data is undesirable. Fine-grained access control is required to provide confidentiality of data. The access control policy is not only role-based, but highly context-based (or rule-based). For instance, patients may have a trust relationship with their current medics, while disregarding the relationship with their former medics. Additionally, the access control and key management mechanism needs to be secure and efficient. Private or secret keys should be securely stored and protected. Finally, potential side channel leakage of medical data should be prevented. For example, the fact that someone takes an HIV test demonstrates that this person is considered at risk. It is desirable to define rules that protect side information without disrupting normal healthcare.

- Data integrity: Medical data integrity should be guaranteed to assure the correctness of the care process. The integrity of logging/auditing data should be ensured for accountability/auditability.

### B. Architecture driven requirements elicitation

In the architecture driven approach, the employed methodology is depicted in Figure 3, to integrate the security and privacy engineering process into the system development lifecycle. In particular, this methodology applies the STRIDE method \([5]\) for security threat modeling and the LIDDUN method \([6]\) for privacy threat modeling and requirement elicitation.

Primarily, Data Flow Diagram (DFD) (Sec. II-B) is used to identify system assets, depicted in Figure 2. This is followed by determining the corresponding security and privacy threats categories at each system asset, i.e., DFD elements, using Table I. The identified security and privacy threat categories are instantiated in the designed TClouds home healthcare system. This is instrumented using security and privacy threat tree patterns \([5]\), \([6]\) providing an overview of the most common preconditions for each threat.

Next, the threat instantiations that are relevant to each assets in the TClouds home healthcare system are documented as misuse cases, each of which presenting a collection of relevant threat scenarios in the system. Moreover, due to both time and budget constraints, the identified threats need to be evaluated and prioritized using risk assessment. Note that details on the risk-analysis process are not discussed in
In order to establish trust in the cloud infrastructure to support critical applications (such as healthcare systems), security and privacy should be built in to assure trustworthiness of the cloud-based system. To bridge between the home healthcare application and cloud infrastructure, as an initial step, a functional infrastructure is sketched in Figure 1. We choose to employ a commodity cloud (e.g., OpenStack [17]). The cloud infrastructure is virtually split into three layers, namely the Infrastructure layer (or Physical Layer), the Platform Layer (or Virtual Layer), and the Application Layer.

The infrastructure layer of the cloud ensembles main physical components and the interactions, and is consolidated to serve the Platform Layer. This layer is split in two layers. One for the physical hardware infrastructure and the other for the abstraction layer infrastructure as provided to manage the hardware infrastructure in the cloud (OpenStack). The SWIFT component of OpenStack provides database functionality via redundant storage. The NOVA

<table>
<thead>
<tr>
<th>Asset</th>
<th>Security threats</th>
<th>Privacy threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td>Spoofing, Tampering</td>
<td>Lack of awareness, Identifiability, Controllability</td>
</tr>
<tr>
<td>Data Flow</td>
<td>Tampering, Information disclosure, Denial of service</td>
<td>Lack of awareness, Identifiability, Policy and consent non-compliance</td>
</tr>
<tr>
<td>Data Store</td>
<td>Tampering, Information disclosure, Denial of service</td>
<td>Lack of awareness, Identifiability, Policy and consent non-compliance</td>
</tr>
<tr>
<td>Process</td>
<td>Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, Elevation of privileges</td>
<td>Lack of awareness, Identifiability, Policy and consent non-compliance</td>
</tr>
</tbody>
</table>

Table I

The identified security and privacy threats at each system asset.

Table II

Misuse case of Information Disclosure at the Physical Activity Repository (Data Store)

| Summary: Data is exposed to unauthorized parties |
| Asset: Monitored physical activity data of the user |
| - The patient: disclosure of personal (sensitive) data |
| Primary attacker: skilled insider / skilled outsider |
| Basic Information Flow: |
| 1) The attacker is a skilled insider that gains access to the TClouds home healthcare system. |
| 2) The attacker spoofs an authorized user or gain access to the database by bypassing the Back-end (31). |
| 3) The attacker retrieves data files without being authorized. |
| Alternative Information Flow: |
| 1) The attacker is a skilled insider that accesses the storage medium where data are stored (e.g., by stealing the hardware or scanning a backup storage medium). |
| 2) The attacker obtains the confidential data. |
| Trigger: Initiated by attacker, i.e., this can happen at any time. |
| Preconditions: |
| 1) The data store is insufficiently protected by access policies. |
| 2) Accessing the Physical Activity Repository is not monitored or it can be bypassed. |
| 3) Insufficient protection of the data storage medium (e.g., hardware or hidden storage volume). |
| Assumptions: |
| 1) There are no side channels available to the attacker. |
| Prevention capture points: |
| 1) The physical activity data files are protected, e.g., with context-aware fine-grained access control, so that the data usage and access are controlled by patient’s security policies, taking in account patients’ privacy policy and consent. |
| 2) Accessing the repository is controlled by system’s access control and security policies. |
| 3) All data access/modification actions are securely logged. |
| 4) Data storage medium are physically protected. |
| Prevention guarantee: |
| 1) Bypassing the access control mechanism to access the repository is prohibited, e.g., bypassing the repository management application (31) to access the Physical Activity Repository (35) is prohibited. |
| 2) Bypassing the patient’s usage and access policy to access data files are prohibited. |
| 3) It is difficult (impossible) to access the storage medium without the right privilege. |

IV. DISCUSSIONS ON INFRASTRUCTURE INTEGRATION AND MITIGATION TECHNIQUES

A. Proposed infrastructure integration

In order to establish trust in the cloud infrastructure to support critical applications (such as healthcare systems), security and privacy should be built in to assure trustworthiness of the cloud-based system. To bridge between the home healthcare application and cloud infrastructure, as an initial step, a functional infrastructure is sketched in Figure 1. We choose to employ a commodity cloud
component, as the main part of an IaaS system, empowers Virtual Machines. The platform layer represents the cloud’s virtual resources, which are directly accessed by customers of the IaaS cloud type. This layer consists of different components, namely virtual machine, virtual storage, and virtual network.

The application layer facilitates the cloud’s home healthcare applications that are hosted using resources in the platform layer. A Service Oriented Architecture (SOA) is deployed to package application functionality as a set of stand-alone and interoperable services. For instance, this enables the front-end and middle-tier processes to interact with a database by using services provided by the back-end data store management Application. To facilitate this, SOA is introduced at the interface connecting the database, so that both the web portal and the patient’s personal monitoring device interact with the SOAP (Simple Object Access Protocol) interface. The use of a SOAP interface provides a standardized and interoperable interface for service users.

B. Provisional mitigation solutions

Regarding the specific security and privacy challenges for the home healthcare system in the cloud (Sec. III-B), mitigation techniques to build in security and privacy in are discussed this section.

One of the main security and privacy challenges is to facilitate patient-centric and data-centric protections for patient’s personal data, such that not only access to electronic health data but also distribution and usage of such data should be controlled according to patient’s privacy policies and consent.

To address the aforementioned threat, taking the example use case in Sec. III-C, it is important to assume that the cloud service providers are semi-trusted. A provisional mitigation solution is to allow patients to encrypt their data (e.g. physical activity data in this scenario) before uploading the data to the cloud. The protection mechanism needs to ensure that patient’s data can only be decrypted by authorized parties according to the patient’s policy and consent.

This leads to the discussion on which components of the system architecture (ref. Figure 2) should be hosted in the cloud. A solution is to host the middle-tier and back-end processes and the data stores in the cloud. The decision made on whether to host the front-end process in the cloud is application dependent. For instance, a client can be an independent application that processes data and use the cloud as part of the supply chain. Alternatively, a client can be a
thin layer (e.g. Web-interface) that performs minimal data processing. In the home healthcare scenario, the front-end processes (24, 25 and 26) can be considered as Web portals that are hosted in the cloud. Patients (6) (or monitoring devices (8)) can access the Web portals via Web interfaces (or the I/O of (USB) hardware) at the client side that are hosted off the cloud.

The purpose is to enable patients to control the distribution and use of their personal data. As a provisional mitigation solution to this, patient’s data needs to be encrypted before uploading to the cloud. This can be achieved mainly in two steps.

1) Encryption of data $M$: the patient’s health or physical activity data ($M$) is encrypted by the patient with a secret key $K$, using for instance a symmetric key encryption algorithm (such as AES).

$$C_M = E_K(M)$$

2) Distribution and management of the secret key $K$: $K$ is then distributed to authorized parties for decryption.

Essentially, cryptographic techniques shifts the protection of secret data to the protection of secret keys. Therefore, developing secure and efficient key management and revocation solutions is one of the main challenges. We propose to  address key distribution with two alternative strategies.

The first approach is to encrypt the secret key $K$ using Attribute-Based Encryption schemes (ciphertext policy attribute based encryption (CP-ABE)) scheme [8]–[11] under the patient’s privacy policy

$$C_K = E_{A_{BE}}(K)$$

The ciphertext $C_K$ can be accessible by everyone. Decryption is only possible if and only if the attribute set of the secret key satisfies the access policy specified in the ciphertext. In particular, though encrypted data can be downloaded by everyone, only authorized parties from the social domain (e.g. family, friends, or fellow patients) or from the professional domain (e.g. the patient’s GP or psychiatrist) are able to decrypt the data. A number of solutions are proposed to extend the CP-ABE schemes with instantaneous attribute revocation [3]. policies updating without decrypting the encrypted data [10], and cross-domain sharing and multi-user decryption [11].

The other approach is to leverage enterprise rights management (ERM) to distribute the secret key $K$ with licenses and distribute the licenses to list of recipients

$$\text{License}_i \leftarrow (ID_i, PK_i, K, \text{sig}_P, \text{Policy}).$$

The license can be a function of the recipient’s identity $ID_i$ and public key $PK_i$, the signature of the data owner (e.g. patient), and the data owner’s security and privacy policy. The secret key $K$ can be encrypted, for instance, based on the patient’s policy using asymmetric encryption algorithm (such as RSA) with the recipient’s public key $PK_i$, as $C_K = E_{PK_i}(K)$. This ERM-based mechanism enables patients to stay in control of both distribution/access and offline usage of their data. This in turn gives patients (as data owners) more control on medical data sharing, meanwhile enabling privacy awareness of patients on how their data is used. In other words, instead of solely relying on security protection on communication networks and infrastructure layer, no matter where and how the data is distributed and stored, data protection is ensured at end points of the communication. A number of proposals on ERM-based protection schemes are available, including [12], and with the extensions regarding license management and permission evaluation [13] and session mobility (the ability to transfer tasks execution across devices) [14]. [15].

V. CONCLUSIONS

In this paper, we presented one of the recent results that has been achieved from the new research project on Trustworthy Clouds (or TClouds) funded by the European Commission. This paper proposes an application of cloud computing for home healthcare and addresses its security and privacy challenges. First, we set the scene by investigating a number of use case scenarios for home healthcare and proposed architecture of a home healthcare system in the cloud. In addition, we identified the challenges for building security and privacy in the proposed cloud-based home healthcare system. We described a comprehensive methodology for integrating security and privacy engineering process into the development lifecycle of the proposed cloud-based home healthcare system. In particular, both business-logic driven and architecture driven strategies are applied to instrument the threat analysis and requirements elicitation focusing on security and privacy. Moreover, we sketched out a functional infrastructure plan to integrate the proposed home healthcare application architecture into a commodity cloud. Finally, we discussed a mitigation solution, with a focus on facilitating patient- and data-centric protection for electronic health data, and highlighted its main components that allow patients to encrypt their own (health or physical activity) data prior to uploading the data to the cloud. We suggested two promising strategies to address the protection and distribution of the secret (decryption) key, namely applying attribute-based encryption schemes or enterprise rights management technologies. The mitigation solutions discussed in this paper are preliminary. Further work is needed to develop trust protocols, in order to elaborate the proposed mitigation solution and to address the identified security and privacy requirements thoroughly.

A. Future work and open issues

The mitigation solutions discussed in this section are still preliminary and need further investigation. We only address one of the main security and privacy threats in terms of
patient’s data protection. We observe that encryption and distribution of the encrypted keys is insufficient to ensure a secure cloud health care system. More mitigation techniques need to be integrated to ensure the trustworthiness of the cloud-based healthcare system. As a next step, it is necessary to design trust protocols and to develop the aforementioned mitigation ideas further. Moreover, solutions to address key management (the secret key revocation and renewal) and policy updating for the proposed home healthcare system remain a research challenge.

In addition, though it is still debatable whether the Health SP (service provider) should be fully trustworthy or not, one may foresee that it is plausible for patient’s privacy to assume that the Health SP is semi-trustworthy (honest but curious). For example, as long as the Health SP can collect the monitored physical activity data from the patients, analyze them and provide coaching advices to patients, it could be not necessary for the Health SP to know the patient’s identity information. As a consequence, the mitigation solution needs to be adapted to facilitate patient’s anonymity and transaction unlinkability towards health service providers. This can be achieved by, for instance, applying anonymous credentials (such as for basic anonymous attribute verification and single-use [16], delegation [17], or endorsement [18]), to enable patients to obtain credentials from an issuer, and generate anonymous tokens proving that the patients have the necessary attributes in order to receive a service from the Health SP or access the Health SP’s resources anonymously. In particular, these tokens do not reveal any patient’s personal information, and it is infeasible to tell whether two tokens were generated by the same credential.

From the implementation perspective, the proposed infrastructure is still under the development phase, and it will be ready for practical evaluation within the coming months.

ACKNOWLEDGEMENTS

This research has been supported by the TClouds project, which is funded by the European Commission’s Seventh Framework Program (FP7/2007-2013) under grant agreement number ICT-257243.

REFERENCES


