DACAR Platform for eHealth Services Cloud

L. Fan, W. Buchanan, C. Thümmler, O. Lo, A. Khedim, O. Uthmani, A. Lawson
Faculty of Engineering, Computing & Creative Industries
Edinburgh Napier University, Edinburgh, UK
L.Fan@napier.ac.uk

D. Bell
NIHR CLAHRC for Northwest London
Imperial College London
London, UK
D.Bell@imperial.ac.uk

Abstract—The use of digital technologies in providing health care services is collectively known as eHealth. Considerable progress has been made in the development of eHealth services, but concerns over service integration, large scale deployment, and security, integrity and confidentiality of sensitive medical data still need to be addressed. This paper presents a solution proposed by the Data Capture and Auto Identification Reference (DACAR) project to overcoming these challenges. The key contributions of this paper include a Single Point of Contact (SPoC), a novel rule based information sharing policy syntax, and Data Buckets hosted by a scalable and cost-effective Cloud infrastructure. These key components and other system services constitute DACAR’s eHealth platform, which allows the secure capture, storage and consumption of sensitive health care data.

Currently, a prototype of the DACAR platform has been implemented. To assess the viability and performance of the platform, a demonstration application, namely the Early Warning Score (EWS), has been developed and deployed within a private Cloud infrastructure at Edinburgh Napier University. Simulated experimental results show that the end-to-end communication latency of 97.8% of application messages were below 100ms. Hence, the DACAR platform is efficient enough to support the development and integration of time critical eHealth services. A more comprehensive evaluation of the DACAR platform in a real life clinical environment is under development at Chelsea & Westminster Hospital in London.

Index Terms—eHealth, Cloud, Platform as a Service, Single Point of Contact, Information Sharing Policy Syntax, Data Buckets, Security, Confidentiality

I. INTRODUCTION

The use of modern communication infrastructures in medicine, and the ubiquitous provision of health care services, are in general subsumed under the term “eHealth” [1]. It necessitates a total reform and digitalisation of a health care system, including its production, supply and management [2]. Such technical innovations are expected to improve the quality of health care services, while lowering both the capital and operational costs significantly. Hence, the governments of the USA [3], Canada [4], UK [5], Japan [7], Korea [2] and of the European Union [6] are keen to shift their traditional health care services to a new paradigm, and to make eHealth a top priority on their policy agendas.

Though considerable research effort has been made in the literature relating to developing individual eHealth services, there is a lack of an open eHealth services platform which would allow the integration of such services into flexible, reliable, multifunctional and cost-effective eHealth systems, as well as their large-scale deployment and delivery. Furthermore, a key challenge in eHealth is to use captured patient data in multiple forms and contexts, while maintaining strict access rights. It has been pointed out that health care data are often subject to a variety of threats and attacks, and that inconsistency and loss of data have resulted in severe consequences [10]. Therefore, a complete eHealth services platform should also provide mechanisms to reinforce the integrity, security, confidentiality and auditability of sensitive medical data throughout their life cycle [8].

The aim of the Data Capture and Auto Identification Reference (DACAR) project [41] is to develop, implement, validate and disseminate a novel, secure, “in-the-cloud” service platform for capture, storage and consumption of data within a health care domain. The objectives of the project include:

- Development of novel distributed and secure infrastructures based on role and inter-domain security polices.
- Smart device and system integration platform based on novel digital forensic security technology.
- Generic risk assessment strategy for smart device and system integration.
- Clinical evaluation and dissemination.

The remainder of this paper is organised as follows. Firstly, the background of this research is presented in Section II. An overview of the DACAR platform is given in Section III, and then the design of the key system components are discussed in Section IV, including medical data capture (Section IV-A), storage (Section IV-B), and consumption (Section IV-C). The current implementation of the DACAR platform, together with its demonstration applications, are outlined in Section V, followed by preliminary evaluation results. Finally, conclusions and future work are presented in Section VI.

II. RELATED WORK

Zhang et. al. have identified a set of security requirements for eHealth application Clouds and proposed a novel security model in [8]. This model is mainly designed for the sharing of Electronic Health Records (EHR), while the DACAR platform aims to support the development, integration and large scale deployment of a wider range of eHealth services.

Kılıç et. al. have proposed to share EHRs among multiple eHealth communities over a peer-to-peer network [34]. A super-peer is used to represent an eHealth community, which is responsible for routing messages and adapting different
meta data vocabularies used by different communities. This super-peer design is similar to a Single Point of Contact (SPoC) of the DACAR platform, yet a SPoC provides more authentication and authorisation functionalities.

For a patient-centric eHealth platform it is crucial to obtain various patients’ consents in an electronic way. Coiera et al. have identified four levels of e-consent, including general consent, general consent with specific exclusions, general denial and general denial with specific consents [35]. DACAR’s policy syntax is able to express all of the above, as well as service authorisation, service subscription and investigation. Furthermore, Pruski has identified the requirements for an e-consent language to capture specific grantees, operations, purposes and period of validity, and proposed a novel language called e-CRL [36]. DACAR’s policy syntax is as competent as e-CRL, and has been successfully applied to other domains besides health care, such as police and social care [37], [38].

Currently, DACAR does not support legacy medical records that are not based on atomic attributes. Amato et al. have proposed a semantic based methodology to extract and classify atomic units of information from legacy “monolithic” medical documents [39]. This provides a useful complement to address the limitation of the DACAR platform.

III. SYSTEM OVERVIEW

A. System Model

DACAR’s system model consists of the following concepts:

1) Domain: A domain refers to a distinct business area that is administered by a single organisation. A health care application may involve multiple domains, such as hospitals, pharmacies, insurance companies and research institutions. The domains cooperate with each other to form a Circle of Trust (CoT), where each CoT member keeps a registry of trustworthy services provided by other CoT members.

2) User: A user refers to a consumer of an eHealth application, which can be a person or an impersonated service. A user must be a member of at least one domain, which is able to resolve the user’s identity into a specific role.

3) Object: An object refers to any entity that is managed by an eHealth system, such as patients and medical devices. An object is identified by a unique identifier (UID) assigned by its owner domain. An eHealth system should withstand content oriented and contextual privacy attacks, which means that even if an adversary has the capability of disclosing sensitive information from storage or communication channels, the adversary is neither able to find out that the information is associated with which object, nor to link the actual source and the destination of a message [11]. Therefore, opaque object pseudonyms shall be used in place of object UIDs [12].

4) Attribute: An object is described by a set of attributes, which are atomic units of information of primitive data types. For example, a patient object may comprise of a name attribute of string type, and blood pressure and heart rate attributes of float type. Keeping attributes atomic offers two advantages. Firstly, it is convenient to generate complex medical documents, e.g. EHRs, from atomic attributes dynamically. Secondly, it is also flexible to share atomic attributes among domains under the governance of fine-grained information sharing polices. In practice, not only the core value of an attribute needs to be stored, but also a number of relevant meta data, e.g. the unit, capturer, location, time and device used to collect the data. When sufficient meta data are preserved, an eHealth application is able to document medical events occurring in the past, and to reconstruct them accurately at a later time.

5) Service: DACAR adopts a Service Oriented Architecture (SOA) to support the integration of eHealth services for data capture, storage and consumption purposes. A service refers to a course-grained, discoverable software entity that exists as a single instance and interacts with applications and other services through a loosely coupled, message-based communication model [13]. From a technical point of view, SOA captures many of the best practices of previous software architectures, including abstraction, autonomy, testability, loose coupling, reusability and statelessness.

6) Hosting Infrastructure: DACAR considers Cloud computing environment as its primary hosting infrastructure. The term “Cloud computing” became popular in 2007, with more than 20 definitions given in [14]. The characteristics of Cloud computing appear well-suited to meet the demand of eHealth applications, because a Cloud is inherently service oriented, loose coupling and strong fault tolerant [15]. Also, the business model of Cloud computing can significantly reduce the IT expertise and financial resources for small and medium sized participants to embark on large scale eHealth applications.

From a computing resource provision point of view, Cloud systems are broadly divided into three categories: public Clouds, private Clouds and hybrid Clouds [16], [17]. Public Clouds are confronted with more security challenges, and thus it is more difficult to guarantee the security of data stored in a public Cloud [18]. Hence, DACAR uses a private Cloud for data storage, and a hybrid Cloud for hosting service instances.

From a service provision point of view, Cloud systems can be classified into at least four categories: Hardware as a Service (HaaS), Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) [14]. DACAR’s eHealth platform belongs to the PaaS category. Firstly, it uses IaaS to establish a scalable and cost-effective Cloud infrastructure. Secondly, it provides a solution stack to facilitate the development, integration and deployment of eHealth Software as a Service.

B. DACAR PaaS

The DACAR platform provides a stack of software components and services, which address the most common eHealth application requirements, so that a developer does not need to implement an eHealth application from scratch. Typical issues addressed by the DACAR platform include:

- Authentication: cryptographic protocols that allow an entity to prove to a remote end its identity.
- Authorisation: individual- and role-based policies that endow entities with access rights to resources.
• Data Persistence – long-term storage of medical attributes, including their core values and meta data.
• Data Integrity – functions that ensure data are accurate, complete and consistent during any operations.
• Data Confidentiality – mechanisms that assure stored or transmitted data are accessible only to those authorised to have access, yet well protected from possible disclosure.
• Audit Trail – mechanisms that keep track of a chronological sequence of audit records pertaining to internal and external events and their implications.

Figure 1 shows a three-layer architecture of the DACAR platform. At the bottom layer are Security and Confidentiality Mechanisms, which are used to meet the authentication, data integrity and confidentiality requirements. DACAR supports federated identity providers running a range of user authentication protocols, from traditional RADIUS [19] and Kerberos [20], to recent OpenID [21] and U-Prove [22]. In addition, DACAR provides libraries and APIs for application developers to implement secure SOAP [23] services. This allows a number of security functions, e.g. digital signature, integrity checksum, hashing and encryption to be applied to application-specific portion of communication payload.

In the middle layer is the Single Point of Contact (SPoC), which is used to meet the authorisation requirement. A SPoC consists of two parts: a policy repository and a policy engine. The policy repository holds domain ontology, i.e. definitions of identities, roles, operations, services, objects, attributes and access rights. Each SPoC represents a single domain, and multiple SPoCs form a peer-to-peer network that represents a Circle of Trust. Information requests are routed through the P2P network to an appropriate SPoC, which uses its policy engine to check the requester’s identity, role, and grants access rights according to existing rules in the policy repository. A SPoC authorisation is issued in the form of a Service Ticket or a Data Ticket, which are security tokens protected by the SPoC’s digital signature.

On the top layer are four system services:
1) The Data Bucket service offers long-term persistence of attributes and supports the Creation, Reading, Updating and Deletion (CRUD) of attribute values and associated meta data. Each attribute is stored in a single data bucket hosted by a Cloud infrastructure, and its CRUD service endpoint is registered with the SPoC of the attribute owner domain.

Any application service can put/get data to/from a data bucket, as long as it satisfies two conditions. Firstly, the service needs to know the qualified name of the target attribute, which is defined by domain ontology. Secondly, a rule needs to be established in the SPoC’s policy repository to allow the service, or in the case of impersonation, the service invoker’s identity or role, to perform CRUD operations over that attribute. If both conditions are met, the service is able to make a data request to the SPoC, which replies with a Data Ticket, carrying a reference to the CRUD service endpoint, a list of authorised operations, period of validity, and one-off session keys encrypted by the public keys of the requester and the CRUD service respectively. A Data Ticket may also carry data anonymisation and sanitisation instructions for the CRUD service to follow, as required by security policies.

2) The Identity Mapping service resolves user and object identifiers into pseudonyms, and vice versa. To enhance the contextual privacy of an eHealth application, opaque pseudonyms should be used in place of transparent user and object Ids, e.g. 12478c1abd instead of PatientNo.253. Hence, the DACAR platform uses pseudonyms whenever it is possible, and only reveals real identities to authorised individuals, roles and services when it is absolutely needed.

3) The Access Control service enables patients to create, edit and remove information sharing policies about their own attributes. DACAR adopts a patient-centric point of view, and regards a patient as the real owner of his/her medical data. Hence, the rights of access to such data should be defined by the patients themselves to their trust circle. The access control service provides a friendly user interface, so that authenticated users can easily set up policies controlling what personal information is available to whom, and what medical services they would like to subscribe to.

4) The Audit Trail service gathers text-based logs from application services, showing who was the active user, and what operations the user has performed during a given period of time. eHealth applications can benefit from the audit trail service in many different ways. Firstly, a sufficiently detailed audit trail enables the reconstruction of medical events and scenarios. Secondly, it keeps track of changes made to a system, and helps to roll them back when necessary. Thirdly, it provides evidence for digital forensics technologies, such as the Digital-DNA [25], to detect security anomalies and carry out countermeasures automatically. Finally, it facilitates the monitoring and analysis of the usage of computing resources, and thus helps to improve on the scheduling and load-balancing of the underlying Cloud infrastructure.

C. Work Flow

Figure 2 shows a storyboard of DACAR’s work flow. Typically, a user consumes an eHealth service developed on the DACAR platform in five steps:
**Fig. 2.** DACAR application service consumption process

**Step 1 Authentication:** The user logs on from one of the federated identity providers using a user name and a password, or other unique personal biometric information.

**Step 2 Request for a service:** The user’s client software forwards the security credential obtained in Step 1 to a responsible SPoC, together with a service request.

**Step 3 Instantiate the service:** The SPoC checks the user’s identity, resolves it into a role, and matches the service request to existing security policies. In the case that the service is provided by the local domain, the SPoC is able to tell whether the user is allowed to consume this service, and to locate the service endpoint within the Cloud. However, if the service is provided by a foreign domain in the CoT, the SPoC will route the service request to another SPoC over the P2P network. For example, when a clinician needs to make contact with a patient’s relatives in an emergency, he sends a request for a police registry service to the local health care SPoC, which forwards the request to a remote police SPoC.

**Step 4 Authorisation:** If the service request is permitted by corresponding security policies, the SPoC that made the decision creates and signs a Service Ticket. This contains the user’s pseudonym and role (supplied by the user’s local SPoC), a reference to the service endpoint, period of validity, and one-off session keys that enable the user’s client software and a service instance to establish a secure SOAP session. Otherwise, a message is returned to tell the reason for rejection.

**Step 5 Consume the service:** Finally, the user’s client software initiates a secure session using the information provided in the Service Ticket and starts to consume the service. If the service requires CRUD operations over certain attributes, the service itself becomes a consumer of related Data Bucket services. In this case, the service needs to go through Steps 1 to 4 to obtain necessary Data Tickets from a SPoC using the service’s own identity, or the service consumer’s identity and role. In the latter circumstance the service is “impersonated”, and shall use the Service Ticket received from its consumer as a complementary security credential – a Service Ticket carries a user’s pseudonym and role, and is signed by a SPoC authority, as explained in Step 4.

**Fig. 3.** Smart medical handheld devices made by CipherLab UK

**IV. DESIGN**

The main objective of the DACAR project is to support the development and integration of eHealth services for the capture, storage and consumption of sensitive medical data. This section elaborates on DACAR’s system components and approaches in accordance with these three aspects.

**A. Data Capture**

DACAR uses Radio Frequency Identification (RFID) and smart mobile devices for simple, efficient and secure capture of medical data. RFID is a technology that can be used to identify, authenticate, track and trace medical objects, as well as to gather information about them and their environment [26]. Usually, an RFID system consists of a passive or active transponder tag, a reader, a software programme for processing the data collected, and a database for data persistence. Smart mobile devices, e.g. mobile phones, PDAs and tablet PCs can be used as RFID readers. They collect attribute values from RFID tags, and then transmit the data to software programmes or databases via a wireless network. Figure 3 shows the smart mobile devices designed and manufactured by CipherLab UK for the DACAR project. Fingerprint access control guarantees that the devices can only be used by authorised medical staff, and secure SOAP services are used to encrypt sensitive medical data over the communication channel.

**B. Data Storage**

After a considerable amount of medical data are captured, a solution is needed for storing them. The DACAR platform keeps attributes in an atomic format to enhance their reusability and manageability, as discussed in Section III-A. Another
Bucket should be discussed from three perspectives: for a life-critical medical application. The design of the Data availability and reliability, when the attribute being hosted is it is convenient to replicate a data bucket to provide high physically hosted by a single Data Bucket, which is flexible rationale behind this design is that an atomic attribute can be physically hosted by a single Data Bucket, which is flexible to deploy, migrate and redeploy within a Cloud. Furthermore, it is convenient to replicate a data bucket to provide high availability and reliability, when the attribute being hosted is for a life-critical medical application. The design of the Data Bucket should be discussed from three perspectives:

1) Attribute-oriented Architecture: This is a key to understanding DACAR’s data bucket approach. Suppose that an eHealth application needs to record patients’ blood pressure and heart rate data in real time. The DACAR platform groups the data samples by attribute types, instead of individual patients. In other words, all patients’ blood pressure data are uploaded to, and preserved by, a data bucket that is dedicated to the Blood Pressure attribute. Similarly, a separate data bucket is established to hold Heart Rate data samples.

In each data bucket, the way of distinguishing data samples for different patients is through their meta data. Table I outlines the relational schema of a data bucket, where a core data entry stores the value of an attribute, and a foreign key refers to a meta data entry that supplements the unit of that value, its owner object, and contextual information such as who captured the value, using which device, at what location and at what time. In fact, each meta data column offers a unique dimension for a data mining application to analyse the data.

2) I/O Operations: Reading and writing of data are carried out via the CRUD service, of which the service endpoint is registered with the SPoC that represents the owner domain of the corresponding attribute. Considering that the cost of storage capacity is becoming lower and lower, the DACAR platform deprecates conventional Update and Delete operations. Instead, it preserves all historical attribute values for audit trail purposes, and returns the most recent value of an attribute, where a single valid value of that attribute is required.

A challenge for the design of the CRUD service is efficient filtering of complex meta data. RESTful web services [24] are proposed to allow a service consumer to send dynamic queries in JSON [27] or AtomPub [28] format to a service provider, to retrieve exactly the data that the consumer needs. This approach can reduce the response time and the cost of network bandwidth and processing power for a data intensive web application. It is especially beneficial to mobile devices having limited computing resources and battery life. However, currently the DACAR platform mainly supports secure SOAP services, which is also the technology used to implement the CRUD service. Until libraries and APIs for secure RESTful services are completed, the DACAR platform currently provides an easy to use query syntax, as outlined by Table II, so that a service consumer can make a dynamic data request over SOAP, without requiring a service provider to implement large numbers of cumbersome web methods.

Suppose that an eHealth application needs to find out the last five body temperature samples, which were above 37.5°C, of critically ill patient JohnDoe at Chelsea & Westminster Hospital, since 9:30am, Jan 1st, 2011. The query below can be used in the case that the application only knows the patient’s transparent Id as Chelwest.CIP.JD0:

```plaintext
[ Object == $Chelwest.CIP.JD0$
  && Value >= 37.5
  && Time >= DateTime(2011,1,9,30,0) ]
[ $SEQ$ > $LAST$ - 5 ]
```

Firstly, the query syntax supports Identifiers, i.e. column names of the data bucket schema, Literal values, and all primitive data Types of the Microsoft .NET framework, as well as the Global Unique Identifier (Guid) and DateTime composite data types. Secondly, it supports common arithmetic, relational and logical Operators. Thirdly, it supports Variables, which are names or key words between two dollar signs.

A variable is needed in the following circumstances:
- Transparent IDs – it is likely that a service consumer only knows the transparent ID of a target object, whereas the object is referred to using an opaque pseudonym in security policies and data buckets. In this case, the service consumer can remind a CRUD service to resolve the ID into a pseudonym using the Identity Mapping Service, by putting the ID between dollar signs.
- Aggregate Values – these are similar to aggregate functions of SQL.
- Sequence Numbers – the CRUD service sorts data samples in an intermediate result set according to their time stamps, and endows each of them with a temporary sequence number, so that a service consumer can specify a sub set that is required from a large collection.

Finally, the query syntax supports a chain of Predicates. A predicate is a partial query between two square brackets, and a chain of predicates serves as multiple “Where” clauses in SQL to narrow down a result set. Also, a predicate provides the scope for a CRUD service to evaluate the actual value of a variable at run time. In other words, the intermediate result set is updated once a predicate is processed, and the actual value of an aggregate or sequence variable is updated accordingly.

3) Data Encryption: A fundamental question is where the data encryption should be performed. Application-level encryption means that sensitive medical data are encrypted and decrypted by application services, which is transparent to the underlying database engine. On the other hand, database-level encryption means that data are encrypted and decrypted by

<table>
<thead>
<tr>
<th>Table</th>
<th>Column</th>
<th>Data Type</th>
<th>Constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Data</td>
<td>Id Value</td>
<td>Integer</td>
<td>Primary Key</td>
</tr>
<tr>
<td></td>
<td></td>
<td>String</td>
<td>Not Null</td>
</tr>
<tr>
<td>Meta Data</td>
<td>Id</td>
<td>Integer</td>
<td>Primary Key</td>
</tr>
<tr>
<td></td>
<td>Unit Object</td>
<td>Guid</td>
<td>Not Null</td>
</tr>
<tr>
<td></td>
<td>Capturer</td>
<td>Guid</td>
<td>Not Null</td>
</tr>
<tr>
<td></td>
<td>Device</td>
<td>Guid</td>
<td>Not Null</td>
</tr>
<tr>
<td></td>
<td>Location</td>
<td>Guid</td>
<td>Not Null</td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>DateTime</td>
<td>Not Null</td>
</tr>
</tbody>
</table>

**TABLE I**

**DESIGN OF THE DATA BUCKET SCHEMA**
the database engine automatically before written to and read from the disk. Both approaches can protect data from storage attacks effectively, e.g. theft of storage media, but they both have disadvantages. Application-level encryption suffer from performance overheads, because the size of encrypted data can be much larger than the original plain text. Furthermore, the encryption process eliminates all data type information, as the original data have to be converted into strings. Comparatively, database-level encryption is more efficient and does not require any change to the application layer. However, it is inherently vulnerable to malicious database administrators.

Currently, the DACAR platform adopts database-level encryption, because sensitive medical data are only stored within a private Cloud, which is set up and maintained by trustworthy administrators. Oracle Transparent Database Encryption [29], or SQL Server Encryption [30] are used to encrypt the data buckets automatically. Of course application developers who plan to store data in a public Cloud can still use application-level encryption schemes such as in [31] and [32].

C. Data Sharing

The most important goal of the DACAR platform is to allow trustworthy individuals, roles, application services and sub-systems to consume captured medical data in many different ways, while maintaining strict access rights. This is achieved by DACAR’s information sharing policy.

The policy syntax is designed as below:

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>[Permission]</strong> indicates the action of the rule and defines whether a request meeting the rule criteria will be permitted or denied.</td>
</tr>
<tr>
<td><strong>[Requester]</strong> identifies the source of a request as a specific individual or the membership of a certain role.</td>
</tr>
<tr>
<td><strong>[Operations]</strong> refer to create, read, update and delete. However, the DACAR platform deprecates update and delete operations, as discussed in the previous section.</td>
</tr>
<tr>
<td><strong>[Attribute]</strong> is a unit of information describing an Object.</td>
</tr>
<tr>
<td><strong>[Object]</strong> refers to any entity that is managed by an eHealth system, such as a patient, a medical equipment, and a software service.</td>
</tr>
<tr>
<td><strong>[Context]</strong> identifies the reason why the information is being shared. It also governs the level of access and permissions associated with information exchange, and thus affects the priority accorded to information requests.</td>
</tr>
<tr>
<td><strong>[Owner]</strong> defines a role with sufficient privileges to manage all aspects of an information element, and to permit or deny access to an information element, as required by legislation and defined responsibilities.</td>
</tr>
<tr>
<td><strong>[Multiplicity]</strong> defines the maximum number of records that can be shared over a period of time.</td>
</tr>
<tr>
<td><strong>[Time Window]</strong> defines the period of validity of a rule using ISO 8601 coordinated universal time format.</td>
</tr>
<tr>
<td><strong>[Compliance]</strong> refers to legislative requirements that affect the exchange of information, as well as data anonymisation and sanitisation instructions.</td>
</tr>
</tbody>
</table>

The policy elements discussed above can be used flexibly while composing security rules for different purposes:

1) **Service Authorisation**: A service authorisation rule allows or denies certain individuals or roles to consume an application service. In this circumstance, the **Object** element is used to identify a service that a rule is about, and **Attributes**, **Context**, **Multiplicity** and **Compliance** elements can be omitted.

2) **Service Subscription**: This represents a patient’s subscription to a specific eHealth service, and thus automatically allows creation and reading of arbitrary numbers of attribute records, as needed by the service to fulfil its functionality. **Context** and **Multiplicity** elements can be omitted in this case.

3) **Specific Consent**: This policy type enables a patient to grant access rights for his/her own attributes to trustworthy individuals and roles in a fine-grained manner. It also allows an impersonated service to access a patient’s information, using its consumer’s identity or role.

4) **General Consent**: Sometimes it is difficult for a patient to name the grantees for a specific consent, because they are unclear, unknown, or hard to describe. A general consent is useful in this situation to facilitate information sharing in a coarse-grained manner. The **Context** element is used for a patient to express the willingness to share his/her information with services for a certain purpose, from a certain domain, or above a certain level of importance.

5) **Investigation**: This policy type is only used in exceptional situations, such as a medical incident investigation, to oblige unconditional information sharing. Hence, elements such as **Requester**, **Context** and **Multiplicity** can be omitted, whereas the **Compliance** element may require the investigator to possess additional security tokens on a per case basis.

Table III gives a summary of DACAR’s policy syntax.

V. IMPLEMENTATION & EVALUATION

Currently, a prototype of the DACAR platform has been implemented using Microsoft .NET 4.0 framework. The SPoC is implemented as a self-hosting WCF service running in Windows Server 2008. A Windows Active Directory domain server is employed as the identity provider that authenticates users using the Kerberos protocol. Data Bucket, Access Control, Identity Mapping and Audit Trail services are all implemented.

<table>
<thead>
<tr>
<th>Query Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Identifiers</strong></td>
<td>Column names, e.g. Value, Unit and Object</td>
</tr>
<tr>
<td><strong>Literals</strong></td>
<td>Literal values, e.g. ‘Alice’ and 3.14</td>
</tr>
<tr>
<td><strong>Types</strong></td>
<td>All C# primitive data types, Guid and DateTime</td>
</tr>
<tr>
<td><strong>Operators</strong></td>
<td>Arithmetic, e.g. +, -, *, / and % Relational, e.g. ==, !=, &gt;, &gt;=, &lt; and &lt;= Logical, e.g. &amp;!,</td>
</tr>
<tr>
<td><strong>Variables</strong></td>
<td>Transparent IDs, e.g. $Chlidren,CIDDS Aggregates values, e.g. SMINS, SMEANs and SMAKS Sequence numbers, e.g. SSEQs, SFIRSTS and SLASTS</td>
</tr>
<tr>
<td><strong>Predicates</strong></td>
<td>Chain of queries, e.g. [prev1][prev2][prev3]...</td>
</tr>
</tbody>
</table>

**TABLE II**

SUMMARY OF DACAR’S DATA BUCKET QUERY SYNTAX
as WCF services, which are hosted by IIS 7 web server. The back end of the Data Bucket is supported by SQL Server 2008, and the front end of the Access Control service is developed using WPF XAML, which can be run on Windows-based PCs and mobile devices. The policy engine is implemented in Java, and both information requests and rules are written in XML.

Furthermore, a proof-of-concept application, namely the Early Warning Score (EWS), has been implemented on top of the DACAR platform. EWS is a medical practice widely used in UK hospitals. The traditional EWS requires medical staff to record and enter six vital signs of a patient on a paper-based observation chart periodically, and to calculate a risk score according to predefined equations. In the case that a patient is evaluated to be “at risk”, the medical staff should make contact with appropriate clinicians. The traditional EWS is prone to mistakes, as the measurement, recording and calculation work all need to be done manually. The new EWS application fully automates this process by capturing vital signs using RFID sensors, transmitting the values to data buckets using smart handheld devices, monitoring patient status constantly in real time, and notifying clinicians by calling or sending messages to their mobile phones. A similar application has been proposed in [9].

Currently, clinical evaluation of EWS is under development at the Chelsea & Westminster Hospital in London. In the mean time, a private Cloud infrastructure has been set up at the Edinburgh Napier University to obtain preliminary experimental results. This Cloud consists of three physical machines, i.e. one Cloud controller running OpenNebular 2.0.1 and two worker nodes running the Xen hypervisor [33]. The hardware specifications of the machines are given in Table IV. Such a configuration is sufficient to host four virtual machine instances for a SPoC server, a Data Bucket server, and two web servers for system and application services respectively. Software has been developed to simulate a number of virtual patients, uploading their vital signs to the data buckets.

Figure 4 shows the end-to-end latency distribution of 1000 application messages. 97.8% of the values were below 100ms, and the rest were between 100ms and 200ms. These results suggest that the DACAR platform imposes only a small delay on application-level messages, and thus is efficient enough to support the development and integration of time critical eHealth services. Although the latency in a real medical set-up might be higher than the simulated result, it should be within the millisecond level, which is still acceptable.

VI. CONCLUSION & FUTURE WORK

This paper presents a novel eHealth services platform designed by the Data Capture and Auto Identification Reference (DACAR) project. Firstly, the DACAR platform facilitates the development of eHealth applications by addressing the most typical requirements, including authentication, authorisation, secure data transmission, persistence, integrity, confidentiality and audit trail. Secondly, it provides a suite of hardware and software solutions to integrate the capture, storage and consumption of sensitive medical data. Thirdly, it supports large-scale deployment and delivery of eHealth services using a scalable and cost-effective Cloud infrastructure.

TABLE III
SUMMARY OF DACAR’S INFORMATION SHARING POLICY SYNTAX

| Policy Elements | Service Authorisation | Service Subscription | Specific Consent | General Consent | Investigation
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Requester</td>
<td>user pseudonym or role</td>
<td>service pseudonym</td>
<td>user pseudonym or role</td>
<td>permit or deny</td>
<td>permit or deny</td>
</tr>
<tr>
<td>Operations</td>
<td>read</td>
<td>create or read</td>
<td>create or read</td>
<td>read</td>
<td>read</td>
</tr>
<tr>
<td>Attributes</td>
<td>n/a</td>
<td>attribute pseudonyms</td>
<td>attribute pseudonyms</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Object</td>
<td>service pseudonym</td>
<td>object pseudonym</td>
<td>object pseudonym</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Context</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>domain &amp; service level</td>
<td>n/a</td>
</tr>
<tr>
<td>Owner</td>
<td>domain pseudonym</td>
<td>domain pseudonym</td>
<td>domain pseudonym</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>n/a</td>
<td>n/a</td>
<td>number of records</td>
<td>number of records</td>
<td>n/a</td>
</tr>
<tr>
<td>Time Window</td>
<td>period of validity</td>
<td>period of validity</td>
<td>period of validity</td>
<td>period of validity</td>
<td>period of validity</td>
</tr>
<tr>
<td>Compliance</td>
<td>n/a</td>
<td>sanitisation instruc</td>
<td>sanitisation instruc</td>
<td>sanitisation instruc</td>
<td>security token</td>
</tr>
</tbody>
</table>

TABLE IV
HARDWARE SPECIFICATIONS OF THE EXPERIMENTAL CLOUD

<table>
<thead>
<tr>
<th>Machine</th>
<th>Cloud Node A &amp; B</th>
<th>Cloud Controller</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Dual Core Xeon 3.0GHz</td>
<td>Quad Core Xeon 2.8GHz</td>
</tr>
<tr>
<td>Memory</td>
<td>32GB DDR3 1,333MHz</td>
<td>12GB DDR3 1,333MHz</td>
</tr>
<tr>
<td>Storage</td>
<td>73GB SCSI 10,000RPM</td>
<td>1TB SATA 7,200RPM</td>
</tr>
<tr>
<td>LAN</td>
<td>Intel Pro 1000PT</td>
<td>Intel Pro 1000PT</td>
</tr>
<tr>
<td>WAN</td>
<td>1Mbps Optical Fibre</td>
<td>1Mbps Optical Fibre</td>
</tr>
</tbody>
</table>

Fig. 4. Application-level message latency distribution analysis
the Data Buckets. This paper outlines the design and implementation of these components, as well as preliminary evaluation results obtained using a demonstration application called the Early Warning Score. The experimental results suggest that the DACAR platform imposes small communication latency on application-level messages, and hence it is sufficiently efficient to support the development and integration of time critical eHealth applications.

In future work, a comprehensive evaluation of the DACAR platform will be carried out in a real medical environment, and the design and implementation of the platform will be improved continuously. Another avenue of future work is to build bridges between DACAR and other commercial eHealth services platforms, such as Microsoft’s Health Vault [40]. The goal is to enable secure sharing of health care information on a larger scale, and ultimately, to support expert-guided proactive patient-centric health care.

ACKNOWLEDGEMENT

This research is partially supported by grant from UK’s Engineering and Physical Sciences Research Council (EPSRC) and Technical Strategy Board (TSB). Also, we acknowledge the hard work of the whole project consortium, including Edinburgh Napier University, Chelsea & Westminster NHS Foundation Trust, Imperial College London, Kodit Database Ltd, GS1 UK Ltd and CipherLab UK Ltd.

REFERENCES