INTERNATIONAL JOURNAL OF COMMUNICATION SYSTEMS
Int. J. Commun. Syst. 2010; **23**:1093–1110
Published online 29 October 2009 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/dac.1073

An extensible and ubiquitous RFID management framework over next-generation network‡

Kideok Cho[1], Sangheon Pack[2], Ted 'Taekyoung' Kwon[1,*,†] and Yanghee Choi[1]

[1]*Multimedia and Mobile Communication Laboratory, School of Computer Science and Engineering, Seoul National University, Seoul 151-742, Korea*
[2]*School of Electrical Engineering, Korea University, Seoul 136-701, Korea*

SUMMARY

Radio frequency identification (RFID) is an enabling technology for a variety of applications in next-generation network (NGN). Although several RFID management frameworks are proposed in the literature, they have limitations in terms of service extension. In this paper, we propose an extensible RFID management framework called *SIP-based RFID management system* (SRMS), which can support various RFID applications. SRMS employs *session initiation protocol* (SIP), which is an Internet standard protocol adopted in NGN for session management and mobility support. SRMS enhances the existing SIP architecture by introducing a *surrogate user agent* (SUA) and an *SRMS name server* (SNS). The SUA performs SIP signaling on behalf of RFID tags with limited capabilities, whereas the SNS provides name resolution services. To illustrate the extensibility of SRMS, we demonstrate two representative application scenarios. Compared with the existing RFID management frameworks, SRMS has advantages of service extensibility and reusability. Copyright © 2009 John Wiley & Sons, Ltd.

Received 31 March 2009; Revised 8 August 2009; Accepted 18 September 2009

KEY WORDS: next-generation network; radio frequency identifier (RFID); session initiation protocol (SIP); IP multimedia subsystem; location management; mobile RFID

1. INTRODUCTION

Next-generation network (NGN) [1] is an advanced and converged network of heterogeneous networks (e.g. Internet and cellular networks), in which all information and services are transported

*Correspondence to: Ted 'Taekyoung' Kwon, Multimedia and Mobile Communication Laboratory, School of Computer Science and Engineering, Seoul National University, Seoul 151-742, Korea.
†E-mail: tkkwon@snu.ac.kr
‡A preliminary version of this paper was presented at the *IEEE International Conference on Pervasive Services* (*ICPS*) *2007*, Istanbul, Turkey, July 2007.

Contract/grant sponsor: MKE/IITA; contract/grant number: 2007-F-038-03
Contract/grant sponsor: MOCIE/IITPA; contract/grant number: RTI05-03-01

based on the Internet Protocol (IP) technology. NGN decouples service functions from network technologies, which allows new services to be created easily independent of the underlying transport technologies. In addition, the following features make NGN the most prominent of the future networks: broadband capabilities, fixed/mobile convergence, unfettered user access to difference service providers, multiple last mile technologies support, and so on. Owing to these characteristics, various applications, such as VoIP, multimedia streaming, instant messaging, presence, and conference applications, can be easily supported in NGN.

On one hand, radio frequency identification (RFID) applications are considered as the emerging applications in NGN. RFID is a technology that allows an object (e.g. product, animal, or person) to be identified from a short distance using radio frequency. The RFID technology performs object identification without the user's intervention, thus becoming one of the representative enabling technologies for ubiquitous computing. Various types of RFID tags and service platforms on a global scale are now available in commercial markets [2].

Representative RFID applications are location tracking in supply chain and mobile RFID services for ubiquitous information retrieval. For location tracking, inexpensive RFID tags are attached to objects to keep track of their current locations. With this location information, the efficiency in business transactions, such as inventory or supply chain management can be significantly improved. On the other hand, in mobile RFID services, it is assumed that RFID readers are embedded in mobile terminals [3]. Hence, users can read the RFID tag information through their mobile terminals and retrieve the information about their interests from application servers.

When designing an RFID framework, extensibility and ubiquity should be considered to support the various RFID applications in NGN. Although several RFID management middlewares are proposed in the literature [4–8], they cannot satisfy these requirements. As most of them define separate network architectures and components, they have limitations in terms of service extension, e.g. interworking with web services and mobile Internet. Moreover, mobile RFID services should be provided at any time regardless of the geographical locations of mobile terminals. However, most of them are specialized middleware solutions only for the RFID data management; they have weaknesses in terms of ubiquity.

Meanwhile, it is advantageous to design RFID management systems based on the IP. As diverse network technologies are expected to be converged based on the IP in NGN, an IP-based RFID management system allows low-cost deployment and easy integration with other IP-based services. Moreover, in NGN, IP multimedia subsystem (IMS) has been adopted to support a variety of multimedia services [9, 10]. IMS employs session initiation protocol (SIP) that provides session control and mobility support [11, 12]. Therefore, we can achieve easy service extension and ubiquitous access in NGN by adopting a SIP-based framework.

In this paper, we propose a novel RFID management framework called *SIP-based RFID management system* (SRMS). SRMS uses SIP as a basic protocol; thus, it can be easily integrated with various Internet applications. In addition, SRMS can be implemented over IMS, which enables ubiquitous mobile RFID services. Our contributions are summarized as follows: (1) we propose an extensible and ubiquitous management framework that can support various RFID applications and can be easily integrated with NGN and (2) we demonstrate two interesting application scenarios to show the extensibility of SRMS.

The remainder of this paper is organized as follows. In Section 2, we summarize the related work on RFID management frameworks. Section 3 describes the SRMS architecture and components. After that, we introduce two applications based on SRMS: location management services and

DOI: 10.1002/dac

ubiquitous information retrieval services in Sections 4 and 5, respectively. In Section 6, we compare the SRMS and the existing RFID management frameworks in terms of extensibility, reusability, and scalability. In Section 7, we analyze the location registration latency in SRMS and present the numerical results. Section 8 concludes this paper with the future work.

2. RELATED WORK

Representative RFID applications are RFID data management (e.g. RFID tag location management in the supply chain or logistics) and mobile RFID services. The well-known RFID management framework is EPCglobal Network [4]. Two main components of EPCglobal Network are *EPC information service* (EPCIS) and *object name service* (ONS). The EPCIS is responsible for RFID tag information processing and data exchanging between EPCISes. On the other hand, the ONS provides a simple lookup service, which takes an RFID identifier (i.e. electronic product code (EPC)) as an input and produces a contact address of the EPCIS managing the EPC of interest as an output [13]. EPCglobal Network can be extended from middleware layer to application layer for interworking with other systems. However, EPCglobal Network does not propose a unified architecture for both RFID data management and mobile RFID services, yet.

Various RFID management frameworks have been introduced in [5–8]. Although some middleware solutions are compliant with the EPCglobal Network [5], others are specialized middleware solutions, and thus have limitations in service extension, e.g. interworking with web services and mobile Internet. In addition, these platforms have limitations in supporting ubiquitous information retrieval services or mobile RFID services: they can only support RFID applications such as inventory, supply chain, or logistics management.

To support mobile RFID services, several systems have been proposed in the literature [14, 15]. In [14], mobile RFID service scenarios and system architectures for each scenario are illustrated. Park *et al.* [15] enhance the security for mobile RFID services. However, they are based on a special middleware or cellular network architectures, and thus have difficulties in supporting various IP-based applications. They only focus on information retrieval using mobile phones and do not consider typical RFID applications such as inventory, supply chain, or logistics management.

In NGN environments, a unified RFID framework for various services enables the easy creation of new RFID services and reduction of management costs. For these reasons, an RFID management framework supporting both EPCglobal Network and mobile RFID services has been introduced in [16]. However, as it integrates both services based on an application level gateway, the gateway can be a single point of failure. In addition, it has protocol translation overhead at the gateway to support both services simultaneously.

For easy integration with NGN and reduction of infrastructure costs, SIP-based location tracking system called SIP-RLTS is proposed in [17]. SIP-RLTS supports push- and pull-based location tracking services. Similar to our framework, SRMS, it adopts the SIP and thus can be integrated with NGN or even with SRMS easily. However, it has the limitation in that it only considers the location management aspect of RFID applications, and not mobile RFID services.

To overcome these shortcomings of previous RFID frameworks and to make RFID applications successful and proliferous, we propose an extensible and ubiquitous RFID management framework over NGN, which supports the various applications within a unified framework.

DOI: 10.1002/dac

3. SRMS: SIP-BASED RFID MANAGEMENT SYSTEM

SRMS exploits session management and mobility support function of SIP for RFID applications. Normally, the hardware capability of an RFID tag is not sufficient to fully support the IP stacks and SIP signaling. In other words, RFID readers are expected to perform only limited functions such as tag reading and collision handling. Therefore, we introduce a surrogate user agent (SUA) in charge of SIP signaling on behalf of RFID tags with limited capabilities. The SUA can be incorporated into the RFID tag or the reader if it has enough capability. In addition, we propose an SRMS name server (SNS), which is designed for the resolution between an RFID identifier (i.e. EPC) and a SIP universal resource identifier (URI). For SIP message routing, SIP servers in the existing SIP architecture can be used.

3.1. SRMS architecture

The SRMS network architecture is depicted in Figure 1. The typical SIP architecture consists of SIP servers and user agents. SIP servers are classified into *proxy*, *redirect*, and *registrar* servers depending on their functions. The proxy server relays received SIP messages to other SIP servers or user agents, whereas the redirect server performs redirection of the received SIP messages [11]. The registrar maintains location information to support location management services, so that it is employed to manage the RFID tags' current locations in SRMS. Each RFID tag has its own home registrar at its home domain. When the RFID tag moves to a foreign domain, it should perform a location registration procedure to inform its home registrar of the up-to-date current location. Therefore, a tracking node (TN) can consult the registrar to find out an RFID tag's location. The application server colocates with the registrar and can support various RFID applications such as health care, transport, logistics, or mobile RFID services. On receiving RFID tag information from the registrar, the application server generates domain-specific RFID events (e.g. baggage shipping events in an airport, a movement of a patient in a hospital, etc) for application decisions. On the other hand, user agents are classified into user agent client (UAC) and user agent server (UAS). The UAC initiates SIP transactions by sending a request message, whereas the UAS responds with reply messages with suitable status codes [11].

3.2. Surrogate user agent (SUA)

The SUA is a SIP user agent that performs SIP signaling on behalf of RFID tags with limited capabilities. The SUA fulfills three functions: *EPC filtering*, *tag-location registration*, and *handling tag-tracking requests*. The EPC filtering performs a similar function to the filtering/collection middleware in EPCglobal Network. As the RFID reader has only a simple scanning function, there may exist multiple readings for one tag. To resolve this redundancy, when an RFID reader scans an RFID tag, it sends scanned information to the SUA. On receiving the information, the SUA first performs two filtering processes: temporal filtering (i.e. whether it is redundant EPC or not) and product filtering (i.e. whether it is EPC of interest or not). For these filterings, an RFID application can register its filtering criteria to SUAs in advance.

After two filtering processes, in case of location management services, the SUA triggers a location registration procedure on behalf of the RFID tag. To do this, the SUA needs a SIP URI for the RFID tag and thus it obtains the SIP URI by referencing the SNS. After obtaining the SIP URI, the SUA updates the current location of the RFID tag by sending a `REGISTER` message

DOI: 10.1002/dac

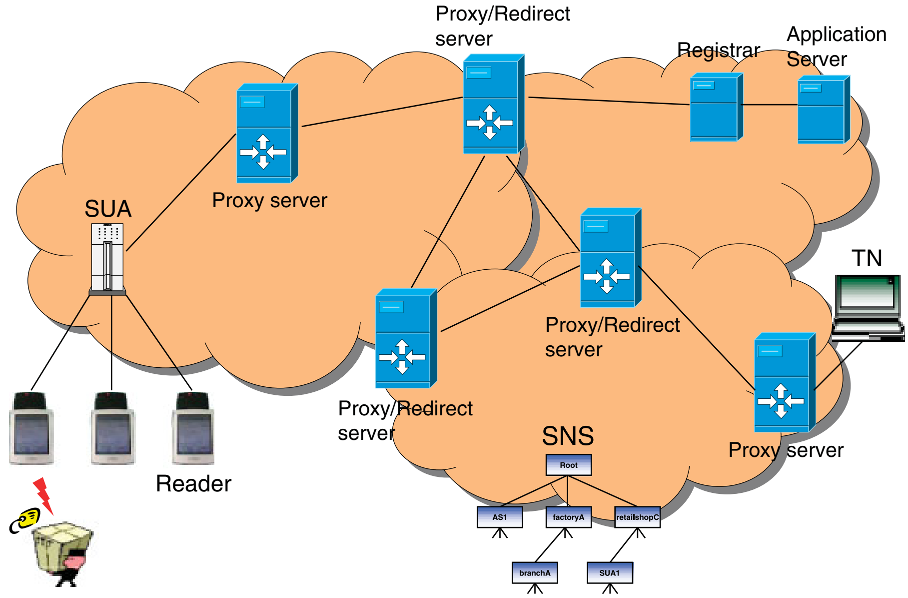

Figure 1. SRMS network architecture.

to the registrar in the tag's home domain. The current location, i.e. the SUA's domain name, is contained in the `contact` field of the `REGISTER` message.

For location tracking of RFID tags, the SUA also plays a crucial role. If a TN wants to track the current location of an RFID tag, the TN sends an `INVITE` message to the corresponding SUA. Then the SUA replies with the tag's current location. In other words, if a tracked tag resides under an SUA's administration, the SUA informs the TN of the tag's current location by sending a `200 OK` message that contains the reader's IP address in the `contact` field. On receiving this `200 OK` message, the TN learns the tag's current location by parsing the `contact` field. In addition, if the TN wants to monitor the tag's condition, it can receive data (e.g. video clip) from the SUA by exploiting the established SIP session. For example, through the established video session, a passenger can check his/her baggages in the airport baggage tracking system.

3.3. *SRMS name server* (*SNS*)

The SNS is a distributed database that maintains mapping between EPCs and SIP URIs. The SNS provides lookup services similar to DNS and ONS; i.e. the SNS takes an EPC as an input and returns a SIP URI. The SUA or TN sends the concerned tag's EPC to the SNS in order to obtain the SIP URI of the tag for location registration, tracking, or mobile RFID services. Then, the SNS finds out the SIP URI for the received EPC and returns it to the SUA or TN.

The EPC is divided into three parts: general manager number, object class, and serial number [18]. As depicted in Figure 2, the general manager number of an EPC is used to generate the home domain in a SIP URI. Namely, the general manager number tells the manufacturer of a product or service provider for mobile RFID and thus it enables the SNS to find out the home domain (e.g. factoryA.com) of the corresponding RFID tag. After that, SNS learns the product type (e.g.

DOI: 10.1002/dac

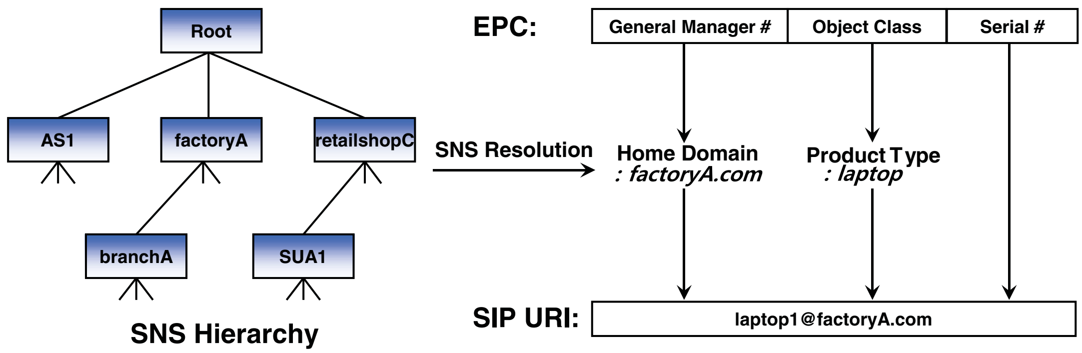

Figure 2. SNS hierarchy and resolution.

laptop) by examining the object class. In addition, the serial number is used to identify an RFID tag uniquely. As a result, the SNS can generate a SIP URI in the form of *laptop1@factoryA.com*.

Owing to the performance and practical reasons, the SNS will be deployed in distributed and hierarchial manners. In this case, when requesting lookup services to the SNS, an SUA or a TN first asks a local SNS (e.g. AS1 in Figure 2). If the local SNS can resolve the request, it replies with a SIP URI. However, if the local SNS has no information about the tag's home domain, it cannot resolve this query; it cannot make a SIP URI. Then, the local SNS asks another SNS recursively, which is at the higher level in the SNS hierarchy, to resolve this query (e.g. Root in Figure 2). When another SNS at a higher level can resolve this query, it delivers the result to the local SNS and the local SNS relays this result to the SUA or the TN. This resolution result can be cached for possible future queries.

4. LOCATION MANAGEMENT SERVICES

In this section, we explain location management services based on SRMS. SRMS makes it possible to track an RFID tag's current location at the resolution of the RFID reader level. That is, a TN can learn the identifier of the RFID reader around which the tag resides. To this end, each RFID reader is assumed to have an IP address. Note that the RFID data management services for health care, transport, or logistics can be supported through the application server colocated with the registrar, in the same way as other RFID middlewares support.

For an illustrative example, it is assumed that the concerned RFID tag's home domain is *factoryA.com* and the tag's SIP URI is *laptop1@factoryA.com*. In addition, SUA1 is an SUA that resides in the foreign domain, *retailShopC.com* and the domain name of SUA1 is *SUA1.retailShopC.com*. The RFID tag moves from *factoryA.com* to *retailShopC.com* and it is identified by RFID reader1 whose IP address is 147.46.0.1 and RFID reader1 is under SUA1's administration.

4.1. Location registration procedure

The location registration procedure is illustrated in Figure 3, where an SRMS entity in domain X is referred to as *SRMS entity* (X).

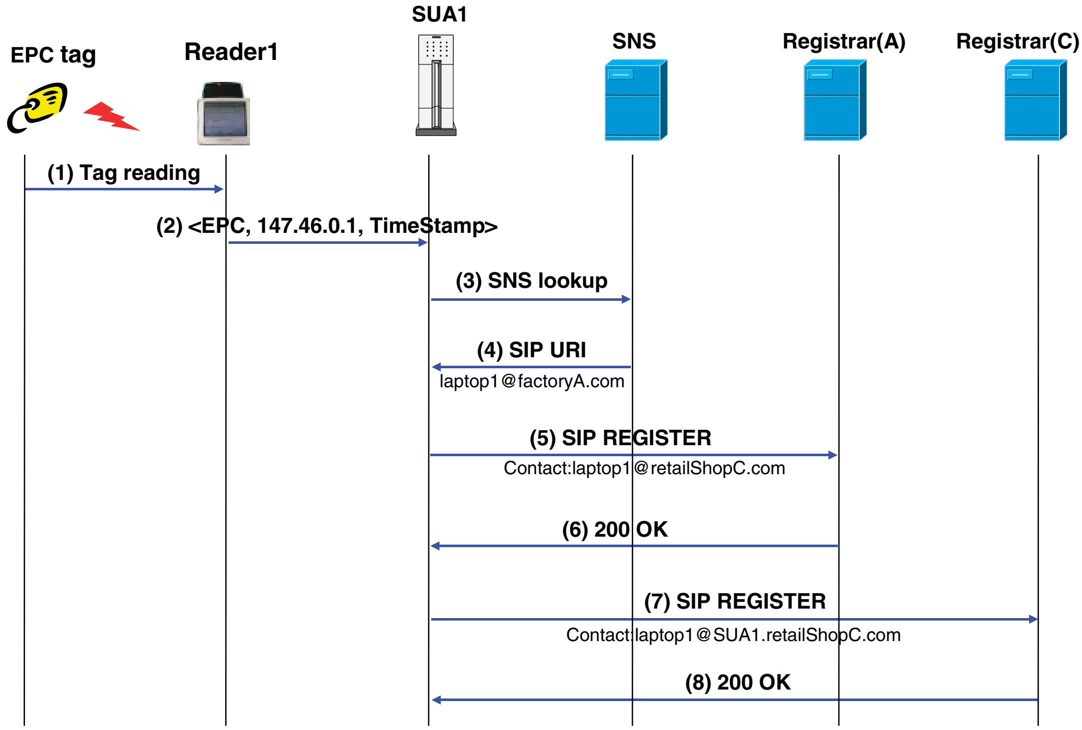

Figure 3. Location registration procedure in SRMS.

Step 1: RFID reader1 scans an RFID tag attached to a laptop and learns the laptop's EPC.

Step 2: The reader1 forwards the read EPC and related information (e.g. its own IP address and time stamp for a filtering purpose) to SUA1. Receiving the EPC-related information, SUA1 performs two filtering processes, where SUA1 checks whether there are redundant readings or not and whether the laptop is an interesting product or not. To reduce the overhead in per-tag processing, multiple readings are buffered during a time interval and they are delivered to the SUA in the form of a batch [19].

Step 3: After completing filtering processes, SUA1 consults the SNS to obtain the SIP URI corresponding to the EPC.

Step 4: The SNS finds out the SIP URI and then returns the SIP URI (i.e. *laptop1@factoryA.com*) to SUA1.

Step 5: SUA1 performs a location registration by sending a `REGISTER` message that contains the current location of the tag, *laptop1@retailShopC.com*, in the `contact` field.

Step 6: If the `REGISTER` message is successfully processed, the registrar (A) sends a response message, `200 OK`, back to SUA1.

Step 7: SUA1 also registers the current location of the tag to its local registrar, i.e. registrar (C) in *retailShopC.com*. The aim of this registration is to forward tracking messages that arrive at *retailShopC.com* domain to the tag's current location, i.e. SUA1. Unlike the previous `REGISTER` message destined for the registrar (A), this `REGISTER` message's `contact` field describes a more detailed tag's location, *laptop1@SUA1.retailShopC.com*. By this message, the registrar (C) learns the precise location of the tag and future SIP messages for the tag can be forwarded to SUA1.

Step 8: The registrar (C) sends a `200 OK` message to SUA1.

As shown in Figure 3, SRMS supports a hierarchical location update [12]. Therefore, SRMS can reduce the amount of SIP signaling similar to Hierarchial Mobile IPv6 [20]. For instance, if an RFID tag moves within the same domain, *retailshopC.com*, a new SUA updates only its local registrar (C), while the home registrar (A) is not updated. For RFID data management, the registrar can forward the received tag information to the colocated application server. The application server processes the received tag information and generates domain-specific RFID events for application decisions. For example, the application server might generate an RFID event saying that 'patient A has physiotherapy in ward B at time $T1$, and has electrotherapy in ward C at time $T2$' for health-care services.

4.2. Location tracking procedure

Figure 4 illustrates the location tracking procedure invoked by a TN in domain B. The following are the detailed location tracking procedures:

Step 1: A TN that would like to know the current location of a laptop first consults the SNS to finds the laptop's SIP URI.

Step 2: The SNS finds out the corresponding SIP URI and returns it to the TN.

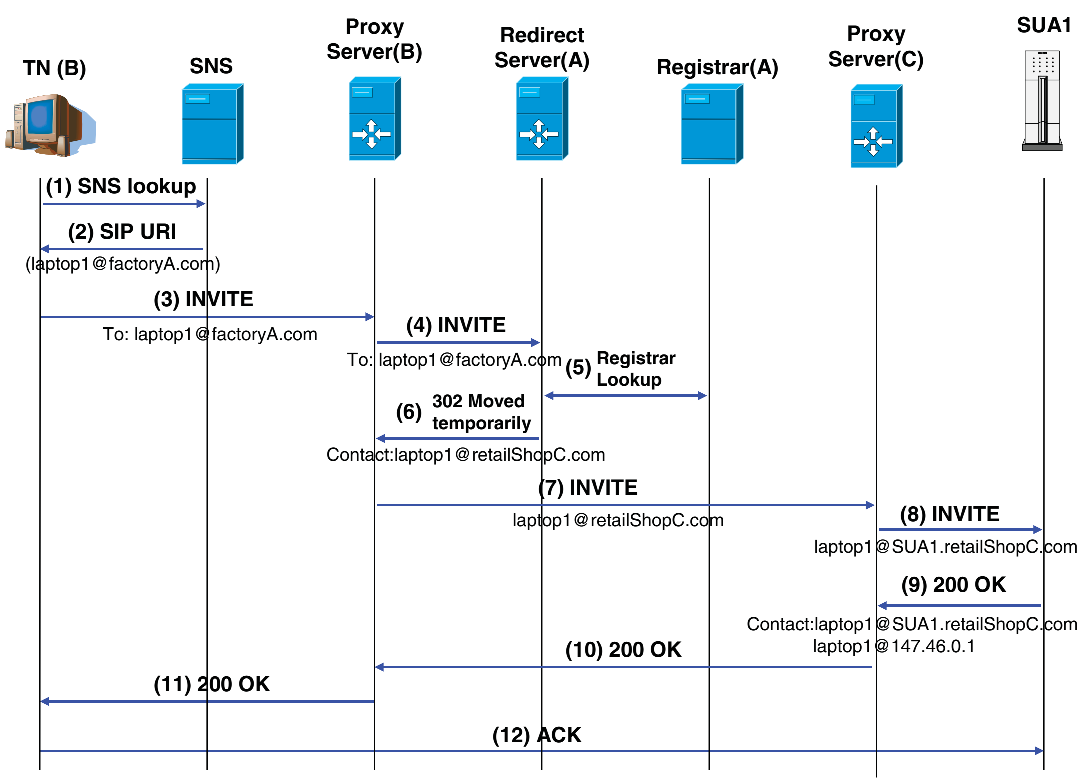

Figure 4. Location tracking procedure in SRMS.

DOI: 10.1002/dac

Step 3: With the found SIP URI, *laptop1@factoryA.com*, the TN sends an `INVITE` message to a proxy server (B) in its domain.

Step 4: The proxy server then forwards the `INVITE` message to the redirect server (A) in the *factoryA.com*.

Step 5: Receiving the `INVITE` message, the redirect server (A) queries the registrar (A) to find the current location of the laptop. For lookup process, directory management protocols such as lightweight directory access protocol [21] can be employed.

Step 6: If the correspondent laptop has moved to another domain, say *retailShopC.com*, the redirect server (A) informs the proxy server (B) of the movement. This notification is accomplished by sending a `302 moved temporarily` response message, which contains the laptop's new SIP URI, *laptop1@retailShopC.com* in `contact` field. This new SIP URI has already been registered through a location registration procedure when the laptop moved to the *retailShopC.com* domain.

Step 7: After receiving the response message, the proxy server (B) sends an `INVITE` message again to proxy server (C) in the *retailShopC.com* domain.

Step 8: By consulting the registrar (C), the proxy server (C) learns that SUA1 manages the RFID reader, which scanned the corresponding RFID tag. Now, the proxy server (C) sends an `INVITE` message to SUA1 by using the learned SUA1's domain name, *SUA1.retailShopC.com*.

Step 9: SUA1 sends back a `200 OK` response message. To track the current location of the laptop at the RFID reader level, the `contact` field of the `200 OK` message includes the IP address of the RFID reader (i.e. 147.46.0.1) as well as the tag's SIP URI (i.e. *laptop1@SUA1.retailShopC.com*).

Steps 10–11: Receiving this response, the proxy server (C) relays the `200 OK` message to the proxy server (B), which in turn relays the response to the TN.

Step 12: At last, the TN sends back the `ACK` message to SUA1.

With the received `200 OK` message, the TN learns the current location of the laptop by parsing the `contact` field in the message. In addition, if the TN wants to monitor the laptop's condition, it can receive data (e.g. video clip) from the SUA1 by exploiting the established SIP session.

5. UBIQUITOUS INFORMATION RETRIEVAL SERVICES

In this section, we describe how to support mobile RFID services via SRMS. Unlike location management services, mobile RFID services provide additional information about an object, i.e. it is a kind of ubiquitous information retrieval service. In mobile RFID service scenarios, objects can be commercial products as well as non-commercial products such as paintings in a museum or product advertisements on the streets.

When users want to acquire related information about products, paintings, or advertisements, they read the RFID tags using their mobile terminals with RFID reader functions. The EPCs of tags are delivered to the application server (or contents server) that maintains additional information about the objects. After receiving these EPCs, the application server responds with the message containing corresponding information back to mobile terminals. After that, mobile terminals display the information to users.

5.1. Mobile RFID service architecture

Mobile RFID services can be supported in SRMS over IMS as shown in Figure 5. We assume that IMS terminals have RFID reader function and SIP signaling capability. That is, the SUA

DOI: 10.1002/dac

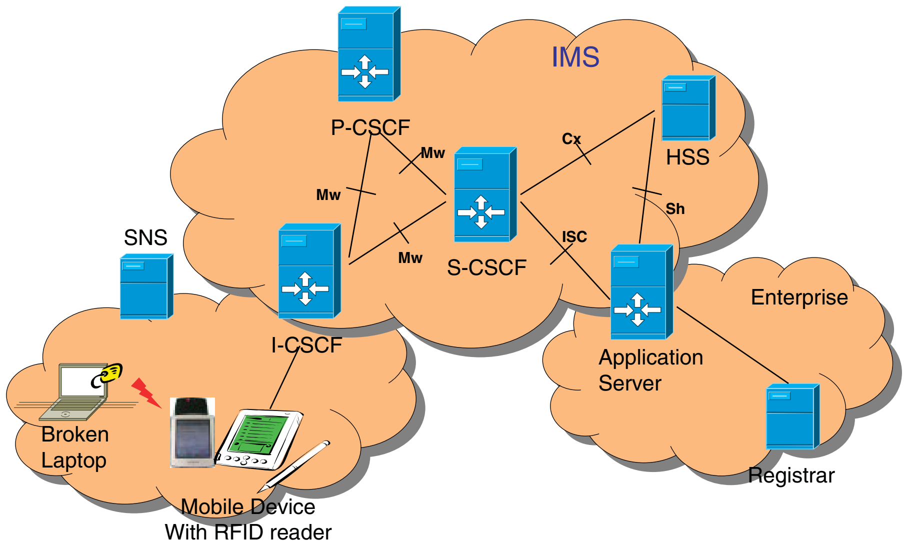

Figure 5. Mobile RFID architecture based on IMS.

functions are incorporated into IMS terminals. In the IMS architecture, there are SIP servers for session control called Call/Session Control Function (CSCF). CSCFs are classified into Proxy (P), Interrogating (I), Serving (S) CSCF depending on their functions. P-CSCF is a kind of proxy server connecting IMS terminals to the IMS network. P-CSCF also performs user authentication and compression/decompression of long SIP messages. I-CSCF also acts as a proxy server. However, unlike P-CSCF, I-CSCF locates at the boundary of the administrative domain. S-CSCF is a central node in IMS SIP signaling. All SIP messages from IMS terminals should be routed through S-CSCF. S-CSCF inspects all SIP messages to decide whether the messages should traverse the application server or not. On the other hand, the Home Subscriber Server (HSS) is a database maintaining user-related information such as location, security, and user profile. In the user profile, there are filter criteria deciding which services are provided to the users. To support mobile RFID services, the filtering criteria should be specified in HSS. According to the predefined filter criteria, the S-CSCF can forward SIP messages to the corresponding application servers.

5.2. *Mobile RFID service procedure*

A ubiquitous information retrieval service scenario is illustrated in Figure 6. In this scenario, we assume that a user wants to obtain repairing information for the broken laptop. The user reads the RFID tag on the laptop and receives movies explaining how to fix the broken laptop from the application server. Thus, we assume that the user's mobile device supports multimedia functions such as audio and video. We also assume that the SIP URI of the EPC for the laptop is *laptopAS@AS1.com*. This SIP URI will be used to route SIP messages to the application server (i.e. AS1.com) to obtain repairing information for the broken laptop. In addition, the username part of the SIP URI (i.e. laptopAS) can tell the application server that it should deliver repairing information for the laptop to the user.

DOI: 10.1002/dac

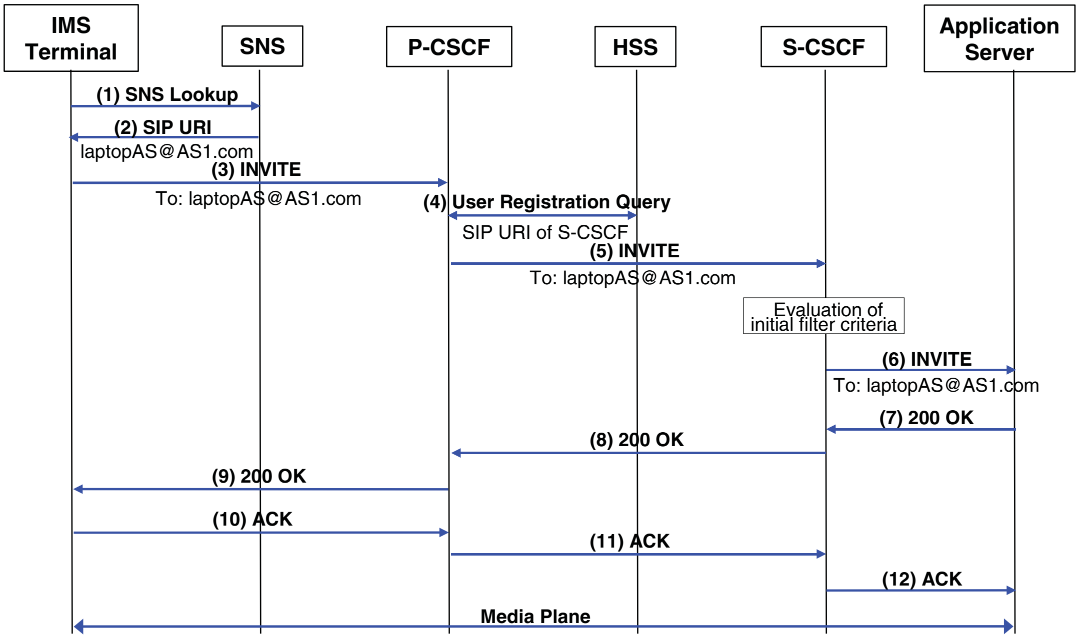

Figure 6. Mobile RFID procedure in SRMS.

Step 1: The user reads the EPC in the RFID tag attached to the broken laptop using a mobile device (i.e. IMS terminal). The IMS terminal consults the SNS to obtain the SIP URI of the application server that maintains repairing information about the laptop.

Step 2: The SNS finds out the SIP URI of the applications server (i.e. *laptopAS@AS1.com*) and returns it to the IMS terminal.

Step 3: The IMS terminal tries to initiate a multimedia session with the application server to receive the repairing information movie. The IMS terminal first sends `INTIVE` to the P-CSCF which is the initial point of contact. This `INVITE` message contains multimedia session information such as IP address, port number, audio, and video codec in the form of Session Description Protocol (SDP) [22].

Step 4: The P-CSCF consults the HSS to see if there exists a dedicated S-CSCF for mobile RFID services. The HSS returns the SIP URI of the dedicated S-CSCF as a response.

Step 5: On receiving the response, the P-CSCF forwards the `INVITE` message, containing the received SIP URI of S-CSCF in `Request-URI` field, to the dedicated S-CSCF.

Step 6: The S-CSCF first evaluates initial filter criteria and learns that the received SIP message is for the mobile RFID services. Then, the S-CSCF forwards the `INVITE` message to the corresponding application server.

Steps 7–9: On receiving the `INVITE` message, the application server sends `200 OK` back to the IMS terminal through S-CSCF and P-CSCF. This `200 OK` message also contains information about a multimedia session in the form of SDP, which will be created between the application server and the IMS terminal.

Steps 10–12: After receiving `200 OK`, the IMS terminal sends an `ACK` message to the application server to create a multimedia session.

DOI: 10.1002/dac

After creating the multimedia session, the application server will send the repairing information movie to the IMS terminal using real-time transport protocol/real-time transport control protocol [23]. The IMS terminal shows the received movie to the user using its own audio and video functions. Or by using instant message service based on Message Session Relay Protocol [24], the IMS terminal informs the user about the URL of the web site maintaining repairing information about the laptop. In this case, the user can access the mobile Internet service using the URL and get the repairing information.

6. DISCUSSIONS

In this section, we compare SRMS with the existing RFID management frameworks in terms of extensibility, reusability, and scalability.

First, SRMS is extensible. The existing RFID management frameworks [6–8] are designed as middleware solutions with special goals. Hence, they have limitations in service extension, e.g. interworking with web services and mobile Internet. In addition, although some frameworks including EPCglobal Network [4, 5] provide extensibility for other services, they are not considering mobile RFID services, yet. On the other hand, SRMS is based on SIP, which uses text-based encoding similar to HTTP, and therefore new services can be easily added at the application level or user level. SRMS can be easily extended and integrated with mobile RFID services and other Internet-based services via SIP. For example, push-based RFID location tracking system [17] can be easily integrated with SRMS without any modification.

Second, SRMS can reuse the existing network infrastructures, e.g. SIP-based VoIP networks. To manage RFID tags, middleware-based RFID management frameworks require a separate network architecture and components [8, 14–16]. On the other hand, as SRMS uses SIP without any modification, SRMS reuses the existing SIP elements such as SIP proxy server, redirect server, and registrar. Currently, SIP is widely employed for a signaling protocol (e.g. VoIP services, IMS) in NGN. Accordingly, it is possible to integrate the SRMS with the existing NGN infrastructures and hence this reusability allows low deployment cost and quick development.

Third, SRMS provides comparable scalability to the EPCglobal Network because it has no central server to manage SIP servers and SNS. SIP servers (i.e. proxy and redirect servers) are distributed in multiple domains and each server performs its own functions in its domain similar to EPCISes. In addition, the SNS can be implemented in a distributed manner as the DNS or ONS does. Similar to the DNS query resolution, EPC lookup in SRMS can be resolved in a local SNS first, and if not, it will be resolved at the higher level in the SNS hierarchy.

7. PERFORMANCE EVALUATION

7.1. Registration latency analysis

To evaluate the performance of SRMS, we model the location registration latency, which is defined from the time instance that a tag is scanned to the time instance that an `INVITE` message for the tag arrives at the home registrar. Figure 7 shows the system model under consideration. In a SIP domain, M readers are managed by an SUA and N SUAs are managed by a local proxy server. Scanned information about X distinct tags is transmitted to the SUA as a unit. Therefore, the SUA

DOI: 10.1002/dac

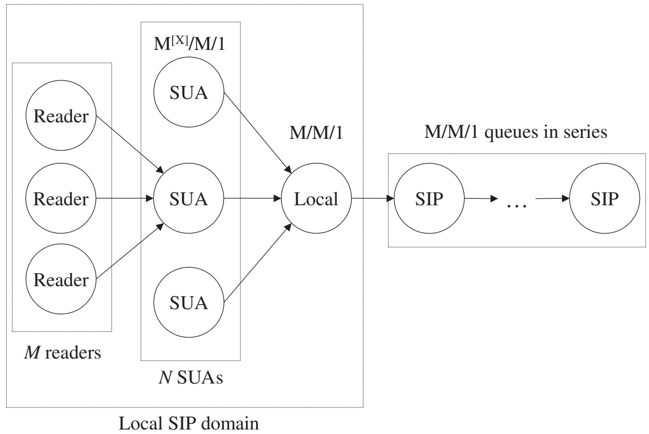

Figure 7. System model.

can be modeled by a queue with batch arrival, i.e. $M^{[X]}/M/1$ queue and it is assumed that X follows a geometric distribution and its probability density function is $f(x)=p(1-p)^{n-1}$ with success probability p. $M^{[X]}/M/1$ queue means Poisson-input, exponential-service, single-server, and the number of tags in any arrival is a random variable X.

For location registration, the SUA sends INVITE messages to the local proxy server. As the departure process of an $M^{[X]}/M/1$ queue can be approximated to a Poisson process, the local SIP server is modeled as an $M/M/1$ queue. On the other hand, intermediate SIP servers and destination registrar are modeled by $M/M/1$ queues in series. As there has been no report about a realistic RFID traffic model, the constant arrival rate or Poisson arrival is used to model the arrival process of RFID tags and $M/M/1$ queueing model for the RFID networks [7, 17]. Similarly, we model the arrival process as the Poisson and the SIP servers as $M/M/1$ queues. However, unlike the previous works, we model the number of tags in arrivals as a random variable as RFID tags usually come in bulk with variable sizes [19].

Now, the location registration latency in SRMS is given by

$$L=t+E[\text{SUA}]+E[\text{Local}]+K\times E[\text{SIP}] \tag{1}$$

where $E[\text{SUA}]$ and $E[\text{Local}]$ are the processing delays at the SUA and local proxy server, respectively, and t is the propagation delay. K is the number of remote SIP severs including intermediate SIP servers and destination registrar. $E[\text{SIP}]$ is the processing delay at the remote SIP server.

Let λ_{SUA} and μ_{SUA} be the arrival and service rates of the SUA $M^{[X]}/M/1$ queue. By [25], the processing latency at the SUA is computed as

$$E[\text{SUA}]=\frac{E[X]+E[X^2]}{2\mu_{\text{SUA}}(1-\rho_{\text{SUA}})E[X]} \tag{2}$$

where $E[X]$ and $E[X^2]$ are the first and second moments of batch size, respectively. ρ_{SUA} is the utilization of the SUA and is given by

$$\rho_{\text{SUA}}=\frac{\lambda_{\text{SUA}}E[X]}{\mu_{\text{SUA}}} \tag{3}$$

DOI: 10.1002/dac

As the local proxy server is modeled as an $M/M/1$ queue, $E[\text{Local}]$ is given by

$$E[\text{Local}] = \frac{1}{\mu_{\text{Local}} - \lambda_{\text{Local}}} \tag{4}$$

where λ_{Local} and μ_{Local} are the arrival and service rates of the local proxy server.

Let $A(s)$ and $B(s)$ be the Laplace transforms of inter-arrival time and service time at the SUA, respectively, and they are given by

$$A(s) = \frac{\lambda_{\text{SUA}}}{s + \lambda_{\text{SUA}}} \quad \text{and} \quad B(s) = \frac{\mu_{\text{SUA}}}{s + \mu_{\text{SUA}}}$$

Then, the Laplace transform of the inter-departure time, $D(s)$, is derived as

$$D(s) = (1 - \rho_{\text{SUA}}) \left(\frac{\lambda_{\text{SUA}}}{s + \lambda_{\text{SUA}}} \right) \left(\frac{\mu_{\text{SUA}}}{s + \mu_{\text{SUA}}} \right) + \rho_{\text{SUA}} \left(\frac{\mu_{\text{SUA}}}{s + \mu_{\text{SUA}}} \right) \tag{5}$$

As ρ_{SUA} is equal to $\lambda_{\text{SUA}} E[X]/\mu_{\text{SUA}}$, $D(s)$ can be rewritten as

$$D(s) = \frac{\lambda_{\text{SUA}}(sE[X] + \mu_{\text{SUA}})}{(s + \mu_{\text{SUA}})(s + \lambda_{\text{SUA}})} \tag{6}$$

Consequently, the average inter-departure time is computed as

$$\begin{aligned} \overline{T} &= (-1) \left. \frac{\mathrm{d}D(s)}{\mathrm{d}s} \right|_{s=0} \\ &= \frac{\mu_{\text{SUA}} + \lambda_{\text{SUA}} - \lambda_{\text{SUA}} E[X]}{\mu_{\text{SUA}} \lambda_{\text{SUA}}} \end{aligned} \tag{7}$$

and the average inter-departure rate is $1/\overline{T}$. Thus, the inter-departure process of an SUA is approximated to an exponential distribution with rate $\mu_{\text{SUA}} \lambda_{\text{SUA}} / (\mu_{\text{SUA}} + \lambda_{\text{SUA}} - \lambda_{\text{SUA}} E[X])$ and a local proxy server covers N SUAs. Therefore, λ_{Local} is given by

$$\lambda_{\text{Local}} = N \times \frac{\mu_{\text{SUA}} \lambda_{\text{SUA}}}{\mu_{\text{SUA}} + \lambda_{\text{SUA}} - \lambda_{\text{SUA}} E[X]} \tag{8}$$

Let $E[S]$ and ρ_{SIP} be the average service time and utilization of a remote SIP server, respectively. Then, the average processing time at the remote SIP server is expressed as

$$E[\text{SIP}] = \frac{E[S]}{1 - \rho_{\text{SIP}}} \tag{9}$$

The location tracking latency can be expressed as a similar equation to the location registration latency except the fact that the tracking latency does not include the process delay at the SUA. We omit the location tracking latency analysis due to the lack of space.

7.2. Numerical results

Table I summarizes the parameter values used for the numerical analysis [26, 27]. The propagation delay t is set to 40 ms as cited in [27]. The service rates of an SUA and a local proxy server are

DOI: 10.1002/dac

Table I. Parameters for numerical analysis.

Parameter	Value
Propagation delay (t)	40 ms
Service rate of the SUA (μ_{SUA})	125 tags/s
Service rate of the local proxy server (μ_{local})	100 tags/s
Average service time of a remote SIP server ($E[S]$)	10 ms
Utilization of a remote SIP server (ρ_{SIP})	0.5
The number of remote SIP servers (K)	5
The number of SUAs (N)	4

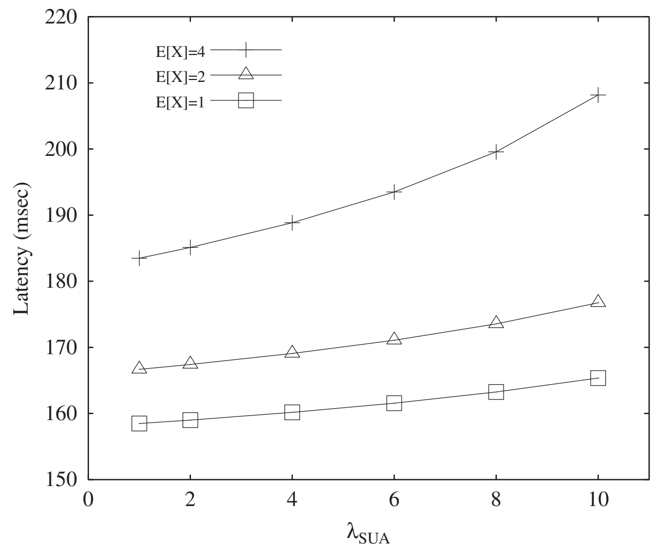

Figure 8. Location registration latency versus λ_{SUA}.

set to 125 tags/s and 100 tags/s, respectively. The average service time of a remote SIP server is assumed as 10 ms and the utilization as 0.5. Furthermore, we assume that there are four SUAs covered by a single local proxy server and a SIP message traverses five remote SIP servers including a registrar.

Figure 8 indicates the location registration latency (L) as a function of the SUA's arrival rate (λ_{SUA}). As the arrival rate of the SUA increases, the registration latency also increases due to the increased processing latency. In addition, when $E[X]$ is large, the SUA should perform more registration procedures, so that the registration latency is also long.

The effects of N and K are illustrated in Figure 9. Intuitively, as N increases, the registration latency increases because more SIP messages should be processed by a local proxy server. Therefore, it is recommended to deploy multiple local proxy servers in a large domain where many SUAs exist. On the other hand, a large K represents that the home registrar is located far away.

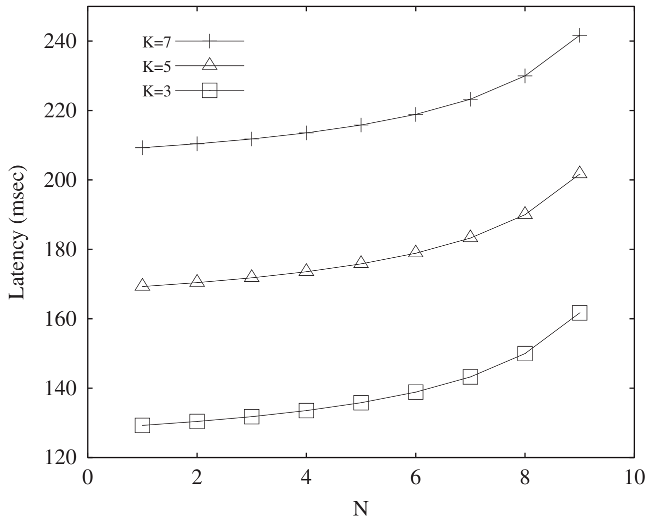

Figure 9. Location registration latency versus N.

Hence, the registration latency increases with the increase in K. This observation indicates that it is possible to reduce the location latency by employing hierarchical registration [12].

8. CONCLUSION

In this paper, we propose an extensible and ubiquitous RFID management framework in NGN called SIP-based RFID management system (SRMS). SRMS is based on a standard IP, SIP, which provides an efficient mobility solution and session control in NGN. In SRMS, the SUA is introduced to handle SIP-related messages and tag filtering. In addition, the SRMS name server (SNS) is proposed to provide global naming services between EPCs and SIP URIs. We also demonstrate two representative application scenarios based on SRMS: location management services and mobile RFID services. Compared with the existing middleware-based RFID management frameworks, SRMS has the advantages of extensibility and reusability because it is based on a distributed architecture using SIP. We are currently building a SRMS testbed using partysip SIP proxy server [28] and we will conduct a comprehensive performance study over the testbed. We will also investigate the interoperbility and security issues in the large-scale SRMS.

REFERENCES

1. ITU-T. General overview of NGN. *ITU-T Recommendation Y.2001*, 2004.
2. Want R. Enabling ubiquitous sensing with RFID. *IEEE Computer* 2004; **37**:84–86.
3. Seidler C. RFID opportunities for mobile telecommunication services. *ITU-T Lighthouse Technical Paper*, 2005.
4. EPCglobal. The EPCglobal architecture framework, version 1.2. *EPCglobal Standard Specification*, 2007.

DOI: 10.1002/dac

5. Cheong T, Kim Y, Lee Y. REMS and RBPTS: ALE-compliant RFID middleware software platform. *Proceedings of the ICACT*, Phoenix Park, Korea, 2006.
6. Bo F, Jin-Tao L, Ping Z, Jun-Bo G, Zhen-Hua D. Study of RFID middleware for distributed large-scale systems. *Proceedings of the ICTTA*, Damascus, Syria, 2006.
7. Zhao Y, Gan O. Distributed design of RFID network for large-scale RFID deployment. *Proceedings of the Fourth IEEE International Conference on Industrial Informatics* (*INDIN'06*), Singapore, 2006.
8. Ting Z, Zhang X, Yuanxin O. A framework of networked RFID system supporting location tracking. *Proceedings of 2nd IEEE/IFIP International Conference in Central Asia on Internet*, Tashkent, Uzbekistan, 2006.
9. ITU-T. *NGN FG Proceedings Part 2. ITU-T*, 2005.
10. Camarillo G, Garcia-Martin M. *The 3G IP Multimedia Subsystem* (*IMS*) *Merging the Internet and the Cellular Worlds*. Wiley: New York, 2004.
11. Rosenberg J *et al.* SIP: session initiation protocol. *IETF RFC 3261*, 2002.
12. Wedlund E, Schulzrinne H. Mobility support using SIP. *Proceedings of the ACM WoWMoM*, Seattle, Washington, U.S.A., 1999.
13. EPCglobal. EPCglobal object name service (ONS) 1.0.1. *Ratified EPCglobal Standards*, 2008.
14. Zhu W, Wang D, Sheng H. Mobile RFID technology for improving m-commerce. *Proceedings of the ICEBE*, Beijing, China, 2005.
15. Park N, Lee J, Kim H, Chung K, Sohn S. A layered approach to design of light-weight middleware systems for mobile RFID security (SMRM: secure mobile RFID middleware system). *Proceedings of the 10th IEEE/IFIP NOMS*, Vancouver, Canada, 2006.
16. Han M, Paik I, Lee B, Hong J. A framework for seamless information retrieval between an EPC network and a mobile RFID network. *Proceedings of the IEEE ICCIT*, Dhaka, Bangladesh, 2006.
17. Li Z, Chu C, Yao W. SIP-RLTS: an RFID location tracking system based on SIP. *Proceedings of the IEEE RFID*, Las Vegas, NV, U.S.A., 2008.
18. EPCglobal. EPC generation 1 tag data standards version 1.4. *EPCglobal Standard Specification*, 2008.
19. EPCglobal. The application level events (ALE) specification, version 1.1.1. *Ratified EPCglobal Standards*, 2009.
20. Soliman H *et al.* Hierarchical mobile IPv6 mobility management (HMIPv6). *IETF RFC 4140*, 2005.
21. Wahl M, Howes T, Kille S. Lightweight directory access protocol (v3). *IETF RFC 2251*, 1997.
22. Handley M *et al.* SDP: session description protocol. *IETF RFC 4566*, 2006.
23. Schulzrinne H, Casner S, Frederick R, Jacobson V. RTP: a transport protocol for real-time applications. *IETF RFC 3550*, 2003.
24. Campbell B, Mahy R, Jennings C. The message session relay protocol. *IETF Internet Draft*, *Work in Progress*, 2006.
25. Gross D, Harris C. *Fundamentals of Queuing Theory*. Wiley: New York, 1985.
26. De P, Basu K, Das S. An ubiquitous architectural framework and protocol for object tracking using RFID tags. *Proceedings of the ACM Mobiquitous*, Boston, MA, U.S.A., 2004.
27. Proceedings of the 53rd Internet Engineering Task Force. Available from: http://www3.ietf.org/proceedings/02mar/slides/ippm-4.pdf [23 March 2009].
28. Partysip SIP Proxy Server. Available from: http://www.partysip.org [23 March 2009].

AUTHORS' BIOGRAPHIES

Kideok Cho received his BS degree (magna cum laude) from the School of Computer Science and Engineering, Seoul National University, in 2004. He is currently working toward the PhD degree at the Multimedia & Mobile Communications Laboratory, Seoul National University. He was a visiting researcher at the Advanced Wireless Networking Laboratory, CUNY in 2007. He is a recipient of the Korea Student Aid Foundation (KOSAF) scholarship from 2009 to 2010. Since 2006, he has been a technical secretary of Future Internet Forum, Korea. His research interests include content networks and wireless sensor/RFID networks.

DOI: 10.1002/dac

Sangheon Pack received the BS (2000, magna cum laude) and PhD (2005) degrees from Seoul National University, both in Computer Engineering. Since March 2007, he has been an assistant professor in the School of Electrical Engineering, Korea University, Korea. From 2005 to 2006, he was a postdoctoral fellow in the Broadband Communications Research (BBCR) Group at the University of Waterloo, Canada. He is a recipient of the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award 2009. From 2002 to 2005, he was a recipient of the Korea Foundation for Advanced Studies (KFAS) Computer Science and Information Technology Scholarship. He received a student travel grant award for the IFIP Personal Wireless Conference (PWC) 2003. He was a visiting researcher to Fraunhofer FOKUS, Germany in 2003. His research interests include mobility management, multimedia transmission, and QoS provision issues in the future wireless/mobile networks.

Ted 'Taekyoung' Kwon has been a faculty member in the School of Computer Science and Engineering, Seoul National University (SNU) since 2004. Before joining SNU, he was a post-doctoral research associate at UCLA and at City University New York (CUNY). He obtained the BS, MS, and PhD degrees from the department of Computer Engineering, SNU, in 1993, 1995, and 2000, respectively. During his graduate program, he was a visiting student at IBM T. J. Watson research center and University of North Texas. His research interests lie in sensor networks, wireless networks, IP mobility, and ubiquitous computing.

Yanghee Choi received the BS from Seoul National University, MS from the Korea Advanced Institute of Science, and a DEng in Computer Science from the École Nationale Superieure des Telecommunications Paris, in 1975, 1977, and 1984, respectively. Since 1991 he has been a professor at the School of Computer Science and Engineering, Seoul National University. He was with the Electronics and Telecommunications Research Institute from 1977 to 1979 and spent the years 1984 to 1991 as the director of the Data Communication Laboratory and the Protocol Engineering Center. He also worked at the Centre National D'Étude des Telecommunications, France, from 1981 to 1984, and at the IBM Thomas J. Watson Research Center from 1988 to 1989 as a visiting scientist. He is now the chair of Future Internet Forum, Korea. He is also the dean of the Graduate School of Convergence Science and Technology and the president of the Advanced Institutes of Convergence Technology, Seoul National University. His research interests lie in the field of future Internet, multimedia communications, and wireless networking.

DOI: 10.1002/dac