A Framework for Mobile Healthcare Applications over Heterogeneous Networks

Amjad Soomro and Ruediger Schmitt
Controls, Communications and Health Informatics
Philips Research North America, Briarcliff Manor, NY 10510
Email: [Amjad.Soomro and Ruediger.Schmitt]@philips.com

***Abstract*—Future healthcare applications will require operation over heterogeneous networks due to mobility of patient, caregiver and provisioning of healthcare services anytime anywhere. In contrast with applications using protocols for specific networking technologies, we propose to include a unified middleware to isolate applications from mobility management, client discovery and transport of multimedia traffic. We propose an all IP-based framework based on SIP protocol for the unified middleware. The architecture is described. We describe how handovers over heterogeneous networks could be implemented in this architecture which provide better QoS and low packet loss during transitions. A laboratory prototype to demonstrate the above concepts is presented.**

I. Introduction

In future, it is expected that healthcare services will be provided not only in the traditional setting of hospitals but management and monitoring of chronic conditions will be commonly done outside hospitals such as in homes. Facilities like assisted living and nursing homes will be closely tied to healthcare providers and clinical systems. The patients will be connected with hospitals' clinical systems to feed medical and personal data as they are being transported to a hospital to reduce potential delays in critical care.

In order to provide ubiquitous access wirelessly, multiple heterogeneous wireless technologies will be used. For example, Wi-Fi is commonly deployed in hospitals and homes, whereas cellular technologies are generally accessible while outdoors. In addition, technologies like Zigbee and Bluetooth are used to provide short range communications. Just as consumer devices, such as smartphones, have evolved to include multiple wireless technologies, healthcare devices will follow the same path. At the same time, the traditional boundaries of coverage for wireless technologies are merging, for example, Wi-Fi is being deployed outdoors while cellular access is expanding to be available within hospitals and homes using technologies like repeaters.

Today, healthcare applications are typically tied to one particular type of wireless access technology. For example, voice calls go through cellular voice network and data flows through IP-based networks. Similarly, SMS text messages are used to provide alerts to mobile healthcare givers, implicitly assuming universal coverage of a cellular technology. However, in areas such as buildings with concrete walls or in dead spots in cellular coverage, alternate technology such as Wi-Fi could be available. Therefore, it is desirable to have an application framework which is agnostic of underlying wireless technologies and which works over heterogeneous networks to cater to expanding healthcare application needs.

In this paper we describe one such framework which is all IP based to provide voice, data, video and messaging services. We make the case for using SIP session management protocol as a building block for this architecture. Then we describe how existing applications map in to this architecture. We follow this by mentioning new capabilities in existing applications which can be realized in this architecture. Also, some new application scenarios and functions enabled by this architecture will be presented. As examples, we describe how functionality like handovers and IP address independent session management in heterogeneous networks could be achieved in this framework. Next, we present an overview of the prototype built in laboratory to demonstrate the operation of the above concepts over cellular and Wi-Fi networks. This is followed by our observations and learnings from the prototype operation. Finally we end this paper with discussions and conclusions.

II. The Current Architecture

Consider the following scenario: An emergency responder provides care to a critically ill patient at his home. He attaches a heart monitor and sends ECG data and other vitals of the patient in real-time to a remote healthcare provider. Concurrently, an audio and video session between the responder and the caregiver on a smart phone or tablet allows the caregiver to interact with the patient visually and provide instructions to the responder. This session could potentially exist on the same patient monitoring device. The patient data is transmitted to and, at the same time, relevant historical medical information including the prescription history of the patient is pulled from the patients EHR and displayed in multiple locations. The responder sends text messages to next of kin to inform them of patient condition and get their directions.

Fig. 1 shows typical communication architecture of the applications in the above scenario. The responder's voice and text messages are carried by cellular infrastructure through protocol stack specific to that cellular provider. The patient's real-time ECG streaming data, remote healthcare provider's video and patient's data flows through data path (shown on the right hand side in Fig. 1) composed of TCP/UDP/IP protocols over wired or wireless mediums.

978-1-61284-696-5/11/$26.00

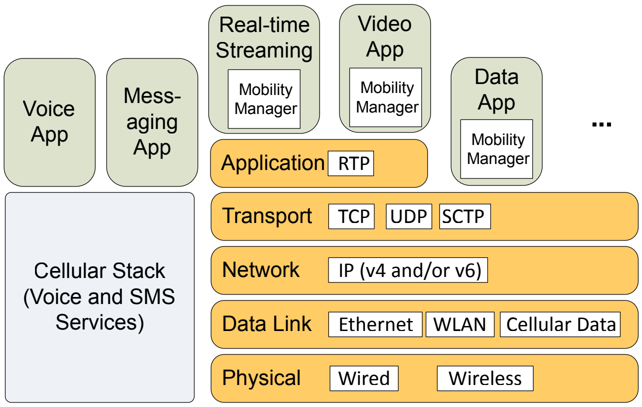

Fig. 1. Current application architecture for applications over heterogeneous networks

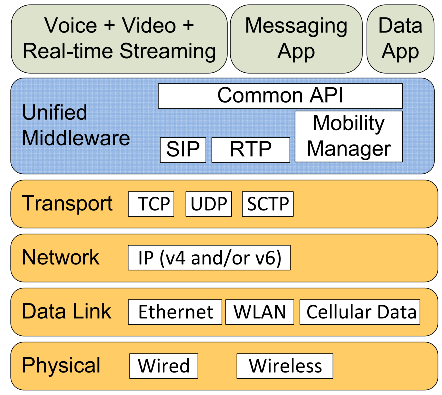

Fig. 2. Proposed application architecture for applications over heterogeneous networks

The above architecture requires availability of both cellular and data paths at any location which may not be the case everywhere, for example, within large apartment complexes. Carrying voice and messaging over IP networks requires applications to use a different protocol suite, complicating the design of the applications. Typically, it is the client which is mobile, therefore, the server applications need to know and maintain the mobile client state to use appropriate protocol to initiate calls or to send text messages. Also, as the patient device or the mobile caregiver moves different networks could be encountered, requiring applications to re-establish sessions on new networks.

III. The Proposed Architecture

In contrast with having distinct protocols specific to each underlying network type, we propose an all IP-based framework for mobile healthcare applications to address the above mentioned challenges. Such an architecture greatly reduces the application complexity by simplifying interface layer and grouping the mobility management and session establishment protocols in one intermediate layer. This architecture allows for the design of applications which maintain continuity of sessions during transitions in heterogeneous network handovers; continuity of sessions is required for connection-based applications. Also, this architecture addresses application needs to push information to caregivers such as making voice calls or sending alerts. We do not address portability of web sessions because web applications do not require session continuity at transport layer and they are agnostic of underlying network type.

Fig. 2 shows the proposed application architecture for communications over heterogeneous networks. We propose to include mobility management, session establishment/management/transfer and real-time data transport in one intermediate layer and provide a common API to all applications. The mobility management component includes functions to assess different available networks and connect with one or more of them. The session management functionality is handled by Session Initiation Protocol (SIP) [1] which is a flexible protocol for creating, modifying and terminating sessions independent of the underlying transport protocols used to carry session data. In IETF protocol suite, SIP works in conjunction with Real-Time Transport Protocol (RTP) [2] and Session Description Protocol (SDP) [3].

SIP provides five facilities for establishing and terminating sessions and they are: 1) User location: discovery of the end system to be used for communications; 2) User availability: determination of the readiness of the called party to engage in communications; 3) User capabilities: determination of the media parameters to be used; 4) Session setup: establishment of session parameters at called and the calling party; 5) Session management: modification, transfer and termination of sessions. We chose to use SIP instead of comparable signaling protocol ITU H.323 [4] because of simplicity of SIP and its flexibility to handle multimedia traffic [5].

Since this architecture uses the IETF protocol stack it enables interoperability among devices and across vendors. However, in order to achieve interoperability, the protocol behavior above the SIP layer, such as SDP parameters, would need to be standardized. This architecture could also be used within an organization to simplify development and management of applications.

A. Applications in the Proposed Architecture

In this architecture, the application end-points are addressed by SIP *Uniform Resource Identifier*, URI, which is similar to an email address, that is, it has a username and a hostname. This is in contrast with the architecture shown in Fig. 1 where the endpoints are addressed by a port number and a hostname. Note that the SIP URI is a logical address which is mapped to a physical hostname or IP address through SIP registration procedure, therefore, it disassociates an end-point's identity from its physical location. This facility is useful for push applications, for example, to send alerts to a caregiver or to initiate calls, because it would not be known at the application server where a mobile caregiver might be at a given time. The applications send and receive data from the same SIP user end-points and transmit the data using the same API regardless of the mobility or the network type of the mobile client. In the case of voice and messaging application, this is the case whether the call goes through a wireless cellular data network or it goes through a data oriented network such as Wi-Fi.

1) Handover Functionality: As a mobile client with a healthcare application moves, it may encounter different heterogeneous networks. Typically, in inter-network transitions, IP address assigned by each network is different; this can also be true for intra-network transitions when connection to a network is reconnected after a brief dropout. Connection oriented applications drop the session during this transition, however, in the architecture presented it is possible to design applications which maintain session across these transitions or IP address changes. In this architecture sessions are established between two SIP URIs and a particular session is referred by a *dialog* identifier. One approach could be to send a SIP re-INVITE message referring to the ongoing dialog with the new IP address. Upon reception of this re-INVITE message the other session end infers that the message is a request to change an existing session rather than create a new one.

Alternatively, during network transitions, unified middleware could create more than one session as SIP allows creation of more than one session between two end-points. Each session could be configured to traverse a different network by using policy based routing. The application data could be forked to go over each session and then re-combined at the other end. During network transitions quality of the data links are dependent on the coverage area of the networks, and in turn, motion of the client. Therefore, quality of data links may not be predicted well in advance. By transmitting over more than one link the probability of dropped packets is reduced by using redundancy of the network during handovers. Some healthcare applications have very low tolerance for the dropped packets and such techniques could be used to improve QoS.

2) Expanded Application and Service Set: By using all IP-based architecture with unified middleware providing a single and consistent point of communicating with the applications, the application code dependency on underlying network type is eliminated. Applications can use any available network, thus expanding the domain where they could be used while simultaneously reducing the coding complexity. Furthermore, since SIP provides built-in registration services, applications which initiate sessions from server can do so in this architecture. For example, hospital initiated video session to a patient or transmission of real-time patient parametric graphs to a mobile caregiver is possible anywhere.

A related issue when communicating with clients behind firewalls and Network Address Translators (NATs) is potential blocking of UDP traffic. IETF has defined protocols for addressing UDP traffic traversal issues through NATs in STUN, TURN and ICE protocols [6]–[8]. These protocols are designed to work in conjunction with SIP. The unified middleware makes use of these protocols to find appropriate ports for use with data traffic.

SIP provides built-in capabilities of instant messaging [9] and these capabilities are exposed to the upper application layers. Since SIP works through data oriented wireless networks such as Wi-Fi and through cellular data networks, message communication from the caregivers can make use of this facility to send messages by using the proposed common framework.

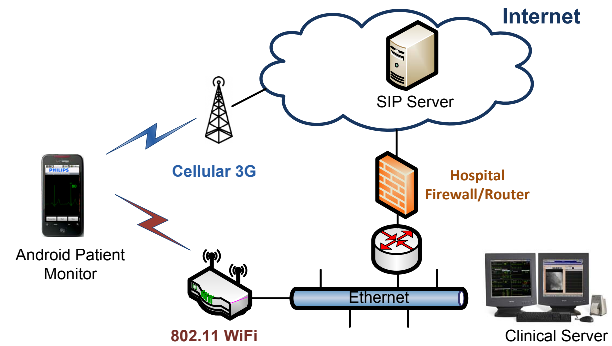

Fig. 3. The laboratory demonstrator for the proposed architecture.

IV. Laboratory Prototype

To validate and demonstrate the concept of the proposed architecture, we built a real-world patient monitoring system shown in Fig. 3. A patient monitoring device sends simulated ECG data to a clinical server via its wireless LAN (WLAN) or cellular data (3G) wireless interfaces. The clinical server, implemented on a Fedora 10 Linux machine, represents a patient monitoring server in a hospital's wired network that displays and further processes the patient ECG data. The hospital's wired network is connected to the Internet through a firewall and a router. We implemented the patient monitoring device on an Android [10] smartphone which has both wireless LAN and cellular data (3G) communication interfaces.

SIP is used to establish two communication sessions between a SIP user agent on the Android patient monitor and a SIP user agent on the clinical server. A FreeSwitch [11] server on a Microsoft Windows machine was used as the SIP proxy and registrar server to facilitate the session establishment. The SIP server has a public IP address so as to be reachable by the SIP user agents. Both WLAN and cellular data interfaces of the Android device were assigned separate private or public IP addresses by DHCP servers in respective networks. We implemented a SIP based communication middleware running on both communication endpoints. The patient monitoring transmitter and receiver applications use this middleware module to exchange data with each other. The emulated patient monitoring application on the Android phone was implemented in Java and the corresponding server application was implemented using PyQt, a Python wrapper around the Qt application framework. The SIP user agents, which are part of the communication middleware on the Android device as well as on the clinical server were implemented using the open source PJSIP and PJMEDIA libraries [12].

Using this demonstrator setup we were able to successfully isolate the emulated patient monitoring applications from the changes in the transport and network layers when moving between WLAN and cellular data connection while maintaining the application session. We found that the effort to build such a system with open source components was reasonable while making the patient monitoring application simpler. We further determined that after establishing network layer connection (as determined by assignment of IP address), the SIP registration and SIP session establishment, over both the cellular data and WLAN networks, can be accomplished in less than one second. The small session re-establishment times reduced data losses during transitions. The cellular data connections were observed to have larger delays and jitter values compared with the ones on WLAN network, though no structured measurements were done.

V. Conclusion

In this paper, we proposed a framework for mobile healthcare applications which works over heterogeneous networks. We proposed to include session establishment and mobility management functions of the communication stack into a middleware. The applications communicate with and address the end-points using SIP URIs. By using SIP URIs, applications are made independent of the underlying network used to carry multimedia traffic. In SIP-based middleware additional benefits are achieved, for example, network handover transparency and improved QoS. At the same time, applications are made simpler by offloading the session establishment and mobility management functionality. We built a laboratory prototype using these concepts for real-time patient monitoring application. Our results confirmed the benefits of the architecture and verified that the continuity of sessions could be maintained across heterogeneous network handover with reduced data losses than the current architecture. Further work on the prototype could be done to test discovery of usable data ports when mobile clients are behind NATs and firewalls. Also, quantitative assessments of variations in delay and jitter of different networks and their effects on patient monitoring application needs to be done.

Acknowledgment

The authors would like to thank Phil Raymond for his comments and suggestions.

References

[1] J. Rosenberg *et al.*, "SIP: session initiation protocol," Internet Engineering Task Force, RFC 3261, Jun. 2002. [Online]. Available: http://www.rfc-editor.org/rfc/rfc3261.txt

[2] H. Schulzrinne *et al.*, "RTP: a transport protocol for Real-Time applications," Internet Engineering Task Force, RFC 3550, Jul. 2003. [Online]. Available: http://www.rfc-editor.org/rfc/rfc3550.txt

[3] M. Handley, V. Jacobson, and C. Perkins, "SDP: Session description protocol," Internet Engineering Task Force, RFC 4566, Jul. 2006. [Online]. Available: http://www.rfc-editor.org/rfc/rfc4566.txt

[4] International Telecommunication Union, "Packet based multimedia communication systems," Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Recommendation H.323, Feb. 1998. [Online]. Available: http://www.itu.int/

[5] H. Schulzrinne and J. Rosenberg, "A comparison of SIP and H.323 for internet telephony," in *Proceedings of International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV 1998)*, Cambridge, England, Jul. 1998, pp. 83–86.

[6] J. Rosenberg *et al.*, "Session Traversal Utilities for NAT (STUN)," Internet Engineering Task Force, RFC 5389, Oct. 2008. [Online]. Available: http://www.rfc-editor.org/rfc/rfc5389.txt

[7] R. Mahy, P. Matthews, and J. Rosenberg, "Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN)," Internet Engineering Task Force, RFC 5766, Apr. 2010. [Online]. Available: http://www.rfc-editor.org/rfc/rfc5766.txt

[8] J. Rosenberg, "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols," Internet Engineering Task Force, RFC 5245, Apr. 2010. [Online]. Available: http://www.rfc-editor.org/rfc/rfc5245.txt

[9] B. Campbell, J. Rosenberg, and Youjian (Eugene) Liu, "Session Initiation Protocol (SIP) Extension for Instant Messaging," Internet Engineering Task Force, RFC 3428, Dec. 2002. [Online]. Available: http://www.rfc-editor.org/rfc/rfc3428.txt

[10] Android software stack for mobile devices. [Online]. Available: http://developer.android.com

[11] FreeSWITCH: A scalable open source cross-platform telephony platform. [Online]. Available: http://www.freeswitch.org

[12] PJSIP and PJMEDIA open source SIP and multimedia libraries. [Online]. Available: http://www.pjsip.org/