Mobile Netw Appl (2007) 12:296–308
DOI 10.1007/s11036-008-0053-5

Interworking Architectures for IP Multimedia Subsystems

Arslan Munir · Vincent W. S. Wong

Published online: 16 April 2008

Abstract The future fourth generation wireless heterogeneous networks aim to integrate various wireless access technologies and to support the IMS (IP multimedia subsystem) sessions. In this paper, we propose the Loosely Coupled Satellite-Cellular-WiMAX-WLAN (LCSCW2) and the Tightly Coupled Satellite-Cellular-WiMAX-WLAN (TCSCW2) interworking architectures. The LCSCW2 and TCSCW2 architectures use the loosely coupling and tightly coupling approach, respectively. Both of them integrate the satellite networks, third generation (3G) wireless networks, worldwide interoperability for microwave access (WiMAX), and wireless local area networks (WLANs). They can support IMS sessions and provide global coverage. The LCSCW2 architecture facilitates independent deployment and traffic engineering of various access networks. The TCSCW2 architecture can provide quality of service (QoS) guarantee. We also propose an analytical model to determine the associate cost for the signaling and data traffic for inter-system communication in these architectures. The cost analysis includes the transmission, processing, and queueing costs at various entities. Numerical results are presented for different arrival rates and session lengths.

This paper was presented in part in *International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine)*, Vancouver, BC, Canada, August 2007.

A. Munir · V. W. S. Wong (✉)
Department of Electrical and Computer Engineering,
The University of British Columbia, Vancouver, Canada
e-mail: vincentw@ece.ubc.ca

Keywords IP multimedia subsystem · network architecture · heterogeneous wireless networks

1 Introduction

The fourth generation (4G) wireless heterogeneous networks are envisioned as the integration of various wireless access technologies such as wireless local area networks (WLANs), third generation (3G) technologies including universal mobile telecommunications system (UMTS) and code division multiple access (CDMA) based CDMA2000 systems, worldwide interoperability for microwave access (WiMAX), and satellite networks. The aim behind this integration is to provide the users with global coverage and to provide the users with the capability to switch between different available access networks (ANs). The success of the future 4G wireless heterogeneous networks depends on the successful integration of the currently available wireless access technologies.

The IP multimedia subsystem (IMS) is standardized by the third generation partnership project (3GPP) and 3GPP2 as a new core network domain [1]. The IMS enables the provision of Internet protocol (IP) based multimedia applications to mobile users, guarantees quality of service (QoS) across different access network technologies and permits service providers to charge according to different policies. In addition, the IMS enables third-party vendors to develop new applications for operators and users. A growing number of telecommunication vendors are beginning to release devices and services based on the IMS [2].

Although various interworking architectures have been proposed before which integrate 3G and WLAN or 3G and satellite networks or 3G and WiMAX individually, to the best of our knowledge, a 4G interworking architecture which can integrate satellite systems, WiMAX, and WLANs with the 3G network has not been considered. In addition, the IMS infrastructure has not been incorporated in the previously proposed interworking architectures. Most of the research related to the performance evaluation of the proposed interworking architectures either analyzes the performance or cost for signaling traffic or data traffic. Little work has been conducted which can evaluate the system performance for both signaling and data. Furthermore, a comprehensive cost analysis which can take into account the transmission, processing and queueing costs of traffic in the interworking architectures has not been considered.

In this paper, we propose two novel 4G interworking architectures namely the Loosely Coupled Satellite-Cellular-WiMAX-WLAN (LCSCW2) architecture and the Tightly Coupled Satellite-Cellular-WiMAX-WLAN (TCSCW2) architecture based on loosely coupled and tightly coupled paradigms, respectively. The IMS has been given special consideration in the proposed interworking architectures. The proposed architectures integrate 3G, WiMAX, WLAN, and satellite networks. We also propose a cost analysis model for the signaling and data traffic for inter-system communication in the proposed interworking architectures. The cost analysis includes the transmission, processing and queueing costs at various network entities. Our analysis takes into account both signaling and data traffic and describes the effect of changing traffic characteristics such as arrival rates and IMS session duration on system cost. The cost analysis will be of significance for the service providers to analyze the individual network elements as well as the proposed architectures comprehensively.

The rest of the paper is organized as follows: The related work is described in Section 2. In Section 3, we present our proposed LCSCW2 and TCSCW2 interworking architectures. In Section 4, we describe an analytical model to determine the cost for IMS signaling and data traffic. Numerical results are presented in Section 5. Conclusions are given in Section 6.

2 Related work

In this section, we summarize various interworking architectures proposed in the literature.

The tightly coupled WLAN-UMTS interworking architecture is proposed in [3–5]. The main concept behind the tightly-coupled approach is to give WLAN appearance of another UMTS access network from the perspective of UMTS core network. In other words, the WLAN is considered like another general packet radio service (GPRS) routing area (RA) in the system. The WLAN-3G interworking function (WIF), which is connected to the serving GPRS support node (SGSN) of the UMTS core network, is responsible for (1) hiding the details of the WLAN from the UMTS core network, and (2) implementing the UMTS protocols for mobility management and authentication which are essential for the UMTS radio access network. For seamless operation, the user equipment (UE) is required to implement the UMTS protocol stack on the top of its standard IEEE 802.11 WLAN network card. Among the disadvantages of tightly-coupled approach is the exposure of the UMTS core network interfaces directly to the WLAN network, which can cause security challenges. Extensive efforts are required for the implementation of WIF especially for the WLAN not owned by the UMTS operators.

The loosely coupled WLAN-UMTS interworking architecture is proposed in [3, 4]. In the loosely coupled architecture, the WLAN connects to the external packet data network and does not have any direct link to 3G network elements such as SGSNs or gateway GPRS support nodes (GGSNs). The loosely coupled architecture has distinct data paths for WLAN and UMTS traffic. The inter-operability with 3G requires the support of mobile IP functionalities or session initiation protocol (SIP) to handle mobility across networks and authentication, authorization, and accounting (AAA) services in the WAG of WLAN. This support is necessary to interwork with the 3G's home network AAA servers. One of the major advantages of the loosely-coupled architecture is that it permits independent deployment and traffic engineering of WLAN and 3G networks. Loose coupling utilizes standard Internet Engineering Task Force (IETF) based protocols for AAA and mobility and therefore, it does not necessitate the introduction of cellular technology into the WLAN network. WLAN uses the extensible authentication protocol (EAP) for authentication of the mobile node by supplying the subscriber identity, subscriber identity module (SIM) based authentication data, and encrypted session keys. In case when WLAN is not owned by the 3G operator and SIM-based authentication is not feasible in the WLAN system, standard user name and password procedures may be deployed. The service continuity during handover to other access networks is not supported efficiently in loose-coupling and causes significant handover latency and packet loss.

A hybrid coupled WLAN-UMTS interworking architecture is proposed in [6]. The hybrid coupled architecture distinguishes the data paths according to the type of traffic and is capable of accommodating traffic from WLAN efficiently with guaranteed QoS and seamless mobility. The tightly-coupled network architecture is chosen for real-time traffic and the loosely-coupled architecture is selected for non-real time and bulky traffic.

A tightly coupled with direct radio access system (TCDRAS) WLAN-UMTS interworking architecture is proposed in [7]. TCDRAS is based on the tight-coupling architecture but it also creates an additional wireless link between the base station of a UMTS cell and WLAN located within the UMTS cell through WLAN to 3G direct controller and transceiver (WGDCT). The signaling in TCDRAS is still routed through the original tightly coupled path for network security purposes.

A loosely coupled with direct radio access system (LCDRAS) WLAN-UMTS interworking architecture is proposed in [8]. LCDRAS is based on the loose-coupling architecture but it also creates an additional wireless link between the base station of a UMTS cell and WLAN located within the UMTS cell through WGDCT. The LCDRAS is capable of dynamically distributing signaling and data traffic to reduce signaling cost and handoff latency via WGDCT.

In addition, an integrated UMTS IMS architecture is presented in [9]. Different scenarios in the 3GPP specifications for WLAN-3G integration are discussed in [4]. The loose coupling and tight coupling interworking architectures are presented in [5]. An architecture which integrates CDMA2000 and 802.11 WLAN is proposed in [10]. The integration of satellite with terrestrial systems is discussed in [11]. In [12], an architecture for UMTS-WiMAX interworking is proposed and the signaling flows for handover from WiMAX to UMTS access network are given. In [13], an architecture is proposed for the integration of WiMAX and UMTS based on loosely coupled approach. In [14], the S-UMTS architecture is proposed and the signaling flows for registration, call handling and handover are given. In [15], an architecture that integrates satellite, WLAN and 3G networks is proposed which requires a third party to handle service level agreements (SLAs).

3 Interworking architectures

In this section, we describe our proposed LCWSW2 and TCSCW2 interworking architectures for IMS.

3.1 The LCSCW2 architecture

Our proposed LCSCW2 interworking architecture is depicted in Fig. 1. This novel interworking architecture integrates satellite networks, 3G wireless cellular networks, WiMAX, and WLANs based on the loosely coupled approach. The areas covered by the satellite spot beams (SSBs), 3G base stations, WiMAX base stations, and WLAN access points (APs) are shown by dotted lines in Fig. 1. Our proposed architecture is compatible with the IMS. Different access networks (ANs) (e.g., 3G networks, satellite networks, WiMAX, and WLANs) can be owned by different service providers (or the same operator). The wireless access gateways (WAGs) of WLAN, WiMAX, satellite and 3G networks are connected to different proxy-call session control function (P-CSCF) servers in IMS via the Internet. In general, each AN has its own separate WAG. In addition, there are separate serving-call session control function (S-CSCF) and interrogating-call session control function (I-CSCF) servers for the two networks. For the establishment of an IMS session between two access networks, the two service providers should have a SLA with each other. IMS networks which are owned by different operators are connected together through an IMS backbone network. The WAG and packet data serving node (PDSN) are connected to the same P-CSCF server if WLAN, WiMAX, satellite and 3G network are owned by the same operator.

The mechanisms involved in the interworking architecture along with the functionalities of various entities are explained below with reference to the 3GPP specification [16]. Access to a locally connected IP network from a WLAN directly is called *WLAN direct IP access*, which is provided by the loosely coupled architecture. The WAG is a gateway via which the data to/from the satellite AN, WiMAX AN, or WLAN AN can be routed to/from an external IP network. In the LCSCW2 architecture, the satellite AN comprises of satellites and satellite fixed earth station (SFES). Satellites convey data and signaling messages between user equipment (UE) and SFES. The SFES performs power control, link control, radio bearer control and paging functions. The SFES is connected to WAG for accessing 3GPP packet switched (PS) and IMS services. The WiMAX AN consists of WiMAX base

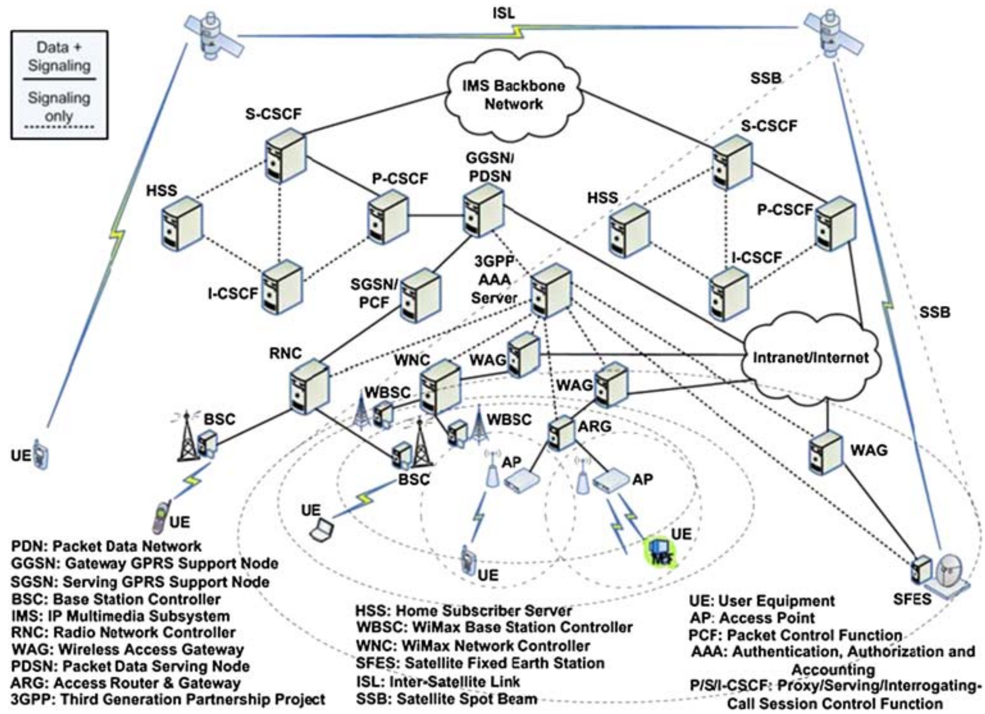

Figure 1 The LCSCW2 interworking architecture

stations, which are controlled by the WiMAX base station controller (WBSC). Several WBSCs are controlled by one WiMAX network controller (WNC). The WNC is connected to WAG to provide WiMAX users with 3GPP PS and IMS services. The LCSCW2 interworking architecture integrates satellite networks, 3G wireless networks, WiMAX, and WLANs based on the loose coupling approach since these ANs connect to the Internet or Intranet via WAG. Then, through the Internet or Intranet, the UE can access CSCF servers of the IMS network.

In the LCSCW2 architecture, the satellite network, WiMAX and WLAN do not have any direct link to 3G network elements such as SGSNs or GGSNs. The LCSCW2 architecture has distinct signaling and data paths for different ANs. The inter-operability with 3G requires the support of mobile-IP functionalities and SIP to handle mobility across networks, and AAA services in the WAG of the AN. This support is necessary to interwork with the 3G's home network AAA servers. The authentication in the ANs is provided through the 3GPP system [16]. The main advantage of the LCSCW2 architecture is that it allows independent deployment and traffic engineering of satellite networks, WiMAX, WLANs, and 3G networks. In addition, this architecture utilizes standard IETF based protocols for AAA and mobility in the WiMAX, WLANs, and satellite networks.

3.2 The TCSCW2 architecture

Our proposed TCSCW2 interworking architecture is depicted in Fig. 2. This interworking architecture integrates satellite networks, 3G wireless cellular networks, WiMAX, and WLANs based on the loosely coupled approach. The areas covered by the satellite spot beams SSBs, 3G base stations, WiMAX base stations, and WLAN access points APs are shown by dotted lines in Fig. 2. Our proposed architecture is compatible with the IMS. Different ANs (e.g., 3G networks, satellite networks, WiMAX, and WLANs) can be owned by different service providers (or the same operator). The packet data gateways (PDGs) of WLAN, WiMAX, satellite and 3G networks are connected to different P-CSCF servers in IMS. In general, each AN has its own separate WAG. In addition, there are separate S-CSCF and I-CSCF servers for the two networks. For the establishment of an IMS session between two access networks, the two service providers should have a SLA with each other. IMS networks which are owned by different operators are connected together through an IMS backbone network. The PDG and PDSN are connected to the same P-CSCF server if WLAN, WiMAX, satellite and 3G network are owned by the same operator.

The mechanisms involved in the interworking architecture along with the functionalities of the important

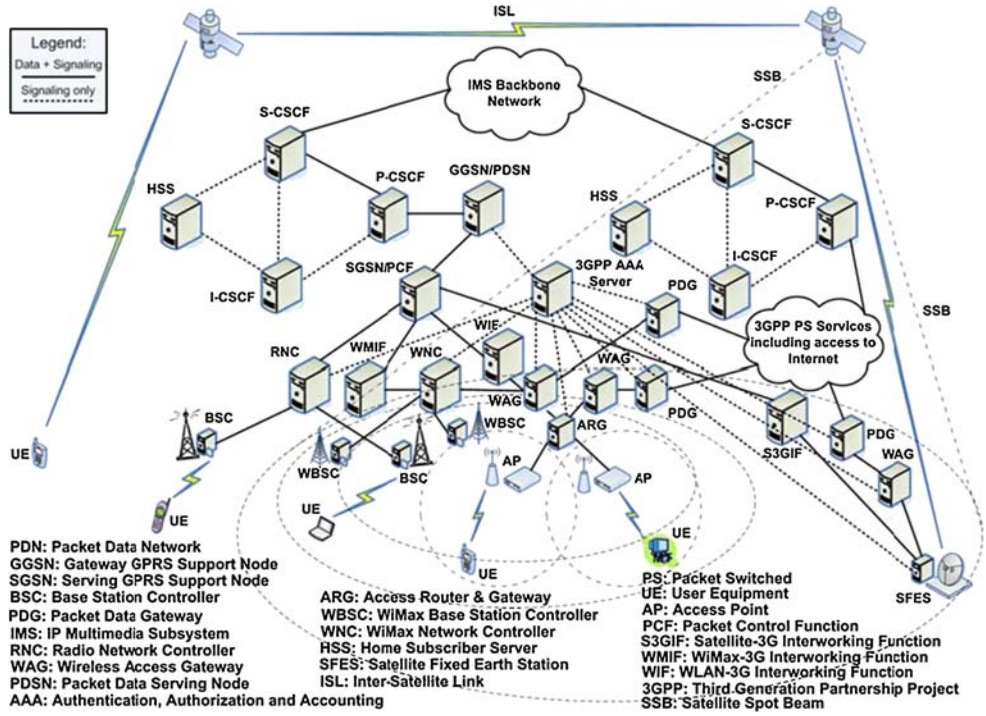

Figure 2 The TCSCW2 interworking architecture

entities are explained below with reference to the 3GPP specification [16]. Access to external IP networks such as IMS, 3G operators network corporate Intranets or the Internet through the 3GPP system is called *WLAN 3GPP IP Access*. The PDG provides WLAN 3GPP IP Access to external IP networks. The WAG is a gateway via which the data to/from the satellite AN, WiMAX AN, or WLAN AN is routed to/from the external IP network. In the TCSCW2 architecture, a UE is identified by multiple IP addresses. For example, in case of a UE in WLAN accessing IMS or 3GPP PS services, the UE is identified by two IP addresses i.e. a local IP address and a remote IP address. A local IP address is used to deliver a packet to the UE in WLAN AN. The local IP address identifies the UE in WLAN AN. The UE's local IP address may be translated by network address translation (NAT) before delivering the packet from UE to other IP network including public land mobile network (PLMN). The remote IP address is used by the data packet encapsulated inside the UE to PDG tunnel. The remote IP address identifies the UE in the network which the WLAN is accessing via PDG. A tunnel is established from the UE to PDG for carrying PS based services traffic in 3GPP IP Access. The data for more than one IP flow and for different services may be carried in one tunnel. It may not be possible to separate individual IP flows and service traffic at intermediate nodes because of the possible encryption of the data including IP header within these tunnels. However, QoS can be assured if the WLAN UE and PDG deploy DiffServ mechanism and appropriately color the differentiated services (DS) field in the external IP header according to the QoS requirement of a particular traffic flow. The PDG assigns remote IP address to the WLAN UE. It registers the WLAN UE's local IP address and binds the UE local IP address with the UE remote IP address. The PDG also performs the encapsulation and decapsulation of packets since it is the terminating/originating point of tunnel between UE and PDG. The WAG performs collection of per tunnel accounting information e.g. byte count, elapsed time etc. and sends this charging information to the 3GPP AAA server [16].

In the TCSCW2 architecture, the WLAN-3G interworking function (WIF), which is connected to the SGSN or PCF of the 3G core network, is responsible for hiding the details of the WLAN from the 3G core network and implementation of the 3G protocols for mobility management, authentication etc. essential for the 3G radio access network. The WIF gives WLAN appearance of another 3G AN from the perspective of 3G core network. In other words, the WLAN is considered like another GPRS Routing Area (RA) in the system. In the TCSCW2 architecture, the satellite access network comprises of satellites and SFES. Satellites convey data and signaling messages exchanged between UE and SFES. SFES performs the power control, link control, radio bearer control and paging functions.

Satellites can be interconnected in an orbit via ISL. SFES is connected to WAG for accessing 3GPP PS and IMS services via PDG. Satellite-3G interworking function (S3GIF) is mainly responsible for connecting satellite systems with core 3G network. The S3GIF, which is connected to the SGSN/PCF of the UMTS/CDMA2000 core network, is responsible for hiding the details of the satellite network from the 3G core network. It is also responsible for the conversion of signaling and packet formats of satellite network to UMTS/CDMA2000 network and vice versa. The WiMAX AN consists of WiMAX base stations controlled by the WBSC. Many WBSCs are controlled by one WNC. The WNC is connected to WAG to provide WiMAX users with 3GPP PS and IMS services via PDG. The WiMAX-3G interworking function (WMIF) connects the WNC to the core 3G network. The TCSCW2 interworking architecture integrate satellite network, 3G, WiMAX, and WLAN based on the tight coupling approach since the satellite network, 3G, WiMAX, and WLAN are directly coupled to the 3G network via interworking functions.

For seamless operation in the TCSCW2 architecture, UEs are required to implement the 3G protocol stack on the top of their standard network cards. Among the disadvantages of tightly-coupled approach is exposure of the 3G core network interfaces directly to the WLAN, WiMAX and satellite network which invites security challenges. Extensive efforts are required for the implementation of interworking functions especially for the ANs not owned by the 3G operators. The 3G core network entities (i.e., SGSN, GGSN) need to be modified to handle the increased load caused by the direct injection of the traffic from other ANs. The TCSCW2 mandates the use of 3G-specific authentication mechanisms based on universal subscriber identity module (USIM) or removable-user identity module (R-UIM) cards for authentication in other ANs. This requires ANs to interconnect to the 3G carriers' signaling system number 7 (SS7) network for performing authentication procedures. Hence, either other ANs interface cards (e.g., IEEE 802.11 WLAN network interface card) be equipped with built-in USIM or R-UIM slots or external USIM or R-UIM cards need to be plugged separately into the UEs. Among the advantages of the TCSCW2 architecture is the possibility of reuse of AAA, mobility management and QoS handling infrastructures of 3G cellular networks. The TCSCW2 architecture enables the provision of 3G services to other ANs users with guaranteed QoS and seamless mobility. However, the 3G core network nodes cannot accommodate the bulky data traffic from the other ANs during busy hours since the core network nodes are designed to support the small-sized data of circuit voice calls or short packets.

4 An analytical model for cost analysis

In this section, we use the inter-system communication cost analysis to evaluate the performance of our proposed LCSCW2 interworking architecture. The cost analysis for the TCWCW2 interworking architecture can be found in [17].

The total inter-system communication cost C_s is given by:

$$C_s = C_t + C_p + C_q \tag{1}$$

where C_t, C_p, and C_q denote the transmission cost, processing cost, and queueing cost, respectively. The transmission cost C_t is the cost incurred due to the transmission of signaling and/or data. It depends on the packet arrival rate, the transmission rate of the link, and the distance between the neighboring network entities. The processing cost C_p is the cost associated with the encapsulation, decapsulation and routing of packets. The queueing cost C_q is the cost incurred due to the queuing of packets in each network entity. Our analysis is applicable to both IPv4 and IPv6 packet types. Also, our analysis is valid for both UMTS and CDMA2000 3G networks.

In the LCSCW2 architecture, communication paths are different when the source node (SN) resides in either WiMAX, WLAN or satellite network. In the following analysis, we assume that SN is communicating via a WLAN and the correspondent node (CN) is using a 3G wireless cellular network. However, the analysis can easily be extended for the case when the SN is communicating via WiMAX or satellite network.

4.1 Available paths for communications

Figures 3 and 4 show the communication path between WLAN and 3G for an IMS session in the LCSCW2 and TCSCW2 architecture, respectively. The solid arrows show the data traffic communication path whereas the dashed arrows indicate the signaling traffic communication path from the SN to the CN. We consider the IMS session establishment signaling [18, 19] where the SN sends a SIP INVITE message to the CN. The SIP 200 OK message is sent from the CN to the SN. The

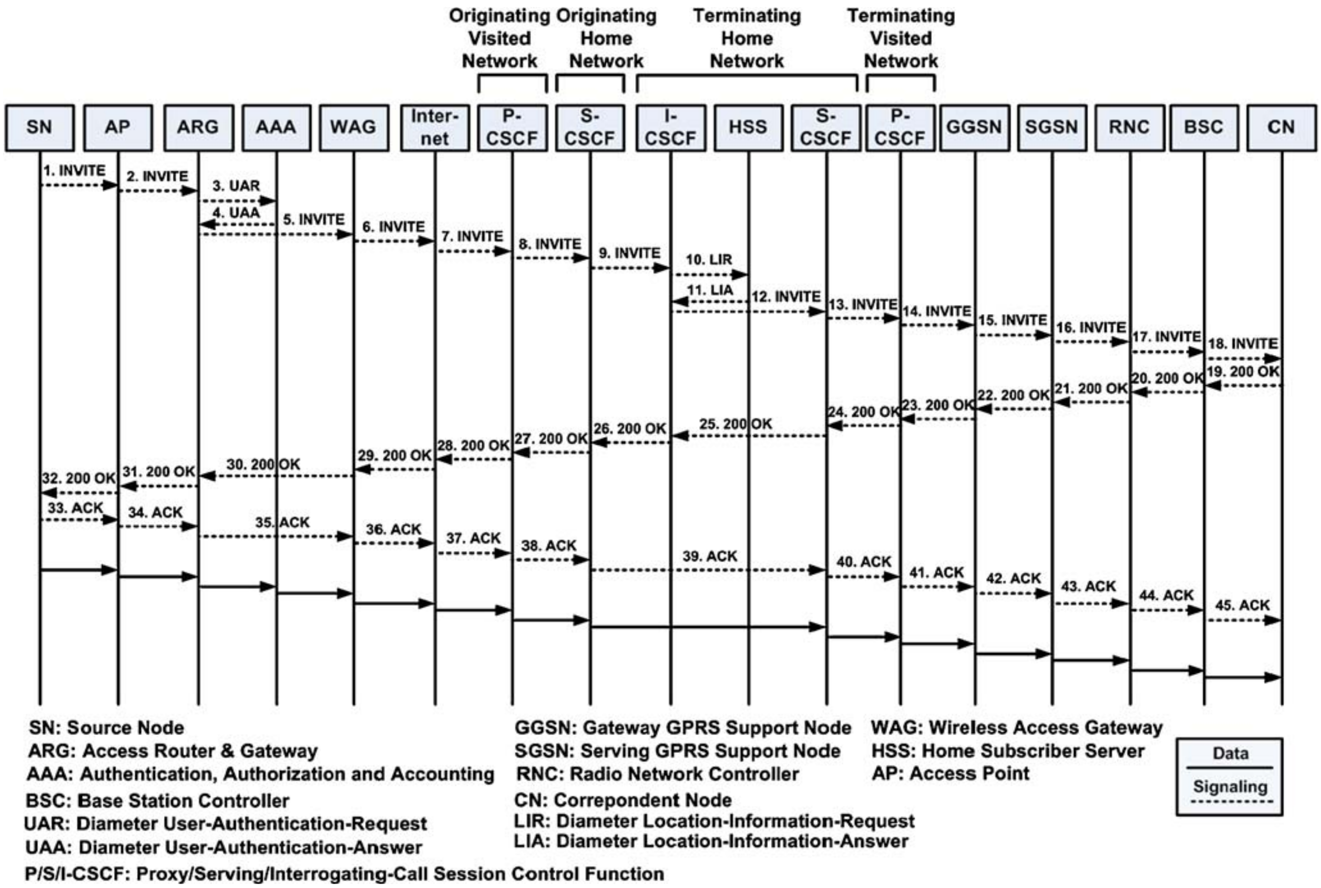

Figure 3 Signaling and data communication paths between WLAN and 3G for IMS session in the LCSCW2 architecture [18, 19]

SIP ACK message from the SN to the CN indicates the completion of session establishment procedure. The signaling incorporates the authentication from the 3GPP AAA server and the query of the user's profile from the HSS database based on Diameter protocol messages [20]. Note that for simplicity, we do not consider the provisional responses such as "100 Trying" in the signaling path.

4.2 Transmission cost

Let λ denote the IMS session arrival rate (requests per second) and $\overline{\ell}$ denote the number of packets per request both for signaling and data from a SN. For transmission and processing cost calculation of IMS signaling traffic, $\overline{\ell}$ is equal to 1 because we assume that one signaling packet can carry one particular signaling message such as 200 OK from one node to its adjacent node. We take into account the traffic coming from other users in the same AN as well as from other ANs by considering background utilization at network entities.

The transmission cost between WLAN and 3G wireless cellular networks for IMS signaling traffic C_t^{sig} is:

$$\begin{aligned} C_t^{sig} = \lambda\overline{\ell}(2\varphi + \psi(3d_{ap-arg} + 2d_{arg-aaa} + 3d_{arg-wag} \\ + 3d_{wag-inet} + 3d_{inet-pcscf} + 6d_{pcscf-scscf} \\ + 4d_{scscf-icscf} + d_{scscf-scscf} + 2d_{icscf-hss} \\ + 3d_{pcscf-ggsn} + 3d_{sgsn-ggsn} + 3d_{sgsn-rnc} \\ + d_{rnc-bsc})) \end{aligned} \quad (2)$$

where φ and ψ are the unit packet transmission costs in wireless and wired link respectively; d_{ap-arg}, $d_{arg-aaa}$, $d_{arg-wag}$, $d_{wag-inet}$, $d_{inet-pcscf}$, $d_{pcscf-scscf}$, $d_{scscf-icscf}$, $d_{scscf-scscf}$, $d_{icscf-hss}$, $d_{pcscf-ggsn}$, $d_{sgsn-ggsn}$, $d_{sgsn-rnc}$, and $d_{rnc-bsc}$ denote the distance between AP and ARG, ARG and AAA, ARG and WAG, WAG and Internet, Internet and P-CSCF, P-CSCF and S-CSCF, S-CSCF and I-CSCF, S-CSCF server of the SN IMS network and the S-CSCF server of the CN IMS network, I-CSCF and HSS, P-CSCF and GGSN, SGSN and GGSN, SGSN and RNC, and RNC and BSC, respectively. The

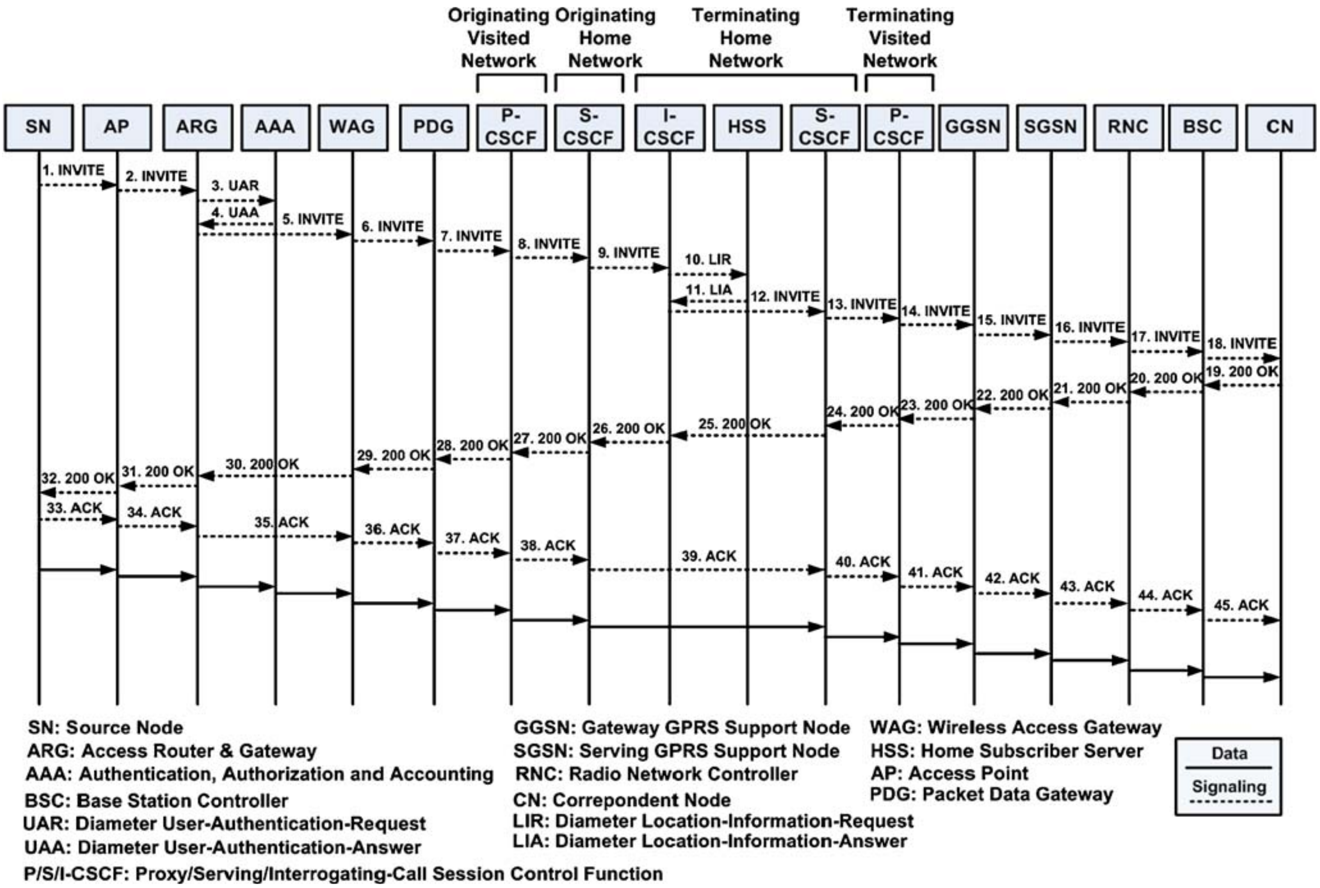

Figure 4 Signaling and data communication paths between WLAN and 3G for IMS session in the TCSCW2 architecture [18, 19]

distance is defined as the number of hops that a packet has traveled.

The transmission cost between WLAN and 3G wireless cellular networks for IMS data traffic C_t^{data} is:

$$C_t^{data} = \lambda\bar{\ell}(2\varphi + \psi(d_{ap-arg} + d_{arg-wag} + d_{wag-inet} + d_{inet-pcscf} + 2d_{pcscf-scscf} + d_{scscf-scscf} + d_{pcscf-ggsn} + d_{ggsn-sgsn} + d_{sgsn-rnc} + d_{rnc-bsc})) \quad (3)$$

By following the same methodology, we can calculate the transmission cost for communication paths between WiMAX and 3G, as well as satellite and 3G for IMS traffic.

4.3 Processing cost

For processing cost calculation, we first assume that N_{bsc} BSCs are connected to each RNC, N_{rnc} RNCs are connected to each SGSN, N_{sgsn} SGSNs are connected to each GGSN, N_{ggsn} GGSNs and N_{pcscf} P-CSCFs are connected to the Internet. In addition, let N_{mn1}, N_{mn2}, N_{mn3}, and N_{mn4} denote the number of users in the coverage area of 3G wireless cellular network, WLAN, WiMAX, and SSB of the satellite network, respectively. The total number of users N in the network can be given as:

$$N = N_{mn1} + N_{mn2} + N_{mn3} + N_{mn4} \quad (4)$$

The processing cost between WLAN and 3G wireless cellular network for IMS signaling traffic C_p^{sig} is:

$$C_p^{sig} = 3C_{p-ap} + 4C_{p-arg} + C_{p-aaa} + 3C_{p-wag} + 3C_{p-inet} + 6C_{p-pcscf} + 6C_{p-scscf} + 3C_{p-icscf} + C_{p-hss} + 3C_{p-ggsn} + 3C_{p-sgsn} + 3C_{p-rnc} + 3C_{p-bsc} \quad (5)$$

where C_{p-ap} represents the processing cost at AP and is given as:

$$C_{p-ap} = \lambda\bar{\ell}\gamma_{ap} \quad (6)$$

where γ_{ap} denotes the unit packet processing cost at AP. The unit packet processing cost includes the cost for encapsulation and decapsulation of packets. Similarly, the terms C_{p-arg}, C_{p-wag}, $C_{p-pcscf}$, $C_{p-scscf}$, and $C_{p-icscf}$ represent the processing costs at ARG, WAG, P-CSCF, S-CSCF and I-CSCF, respectively. Their expressions are similar to that of C_{p-ap} with the only difference that they have their own respective unit packet processing costs. C_{p-aaa} represents the processing cost at the AAA server and is given by:

$$C_{p-aaa} = \lambda\overline{\ell}\left(\gamma_{aaa} + \omega_1\left(\log_{k+1} N + \frac{L}{S}\right)\right) \quad (7)$$

where γ_{aaa} denotes the unit packet processing cost at AAA server. We assume that IP addresses are searched in the lookup table using the multiway and multicolumn search [21]. We also assume that the number of entries in the lookup tables for AAA server and HSS are equal to the total number of users N in the network because 3GPP AAA server based authentication and subscription database HSS are used [16]. In addition, L is the IP address length in bits (e.g. L is 32 for IPv4 and 128 for IPv6), S is the machine word size in bits, and k is a system-dependent constant. In our analysis, ω_i where $i \in \{1, 2, 3, 4\}$ denotes the weighting factors. C_{p-hss} represents the processing cost at HSS and its expression is similar to that of C_{p-aaa} with the only difference that it has its own specific unit packet processing cost. C_{p-ggsn}, C_{p-sgsn}, C_{p-rnc}, and C_{p-bsc} represents the processing costs at GGSN, SGSN, RNC, and BSC respectively with similar expressions as that of C_{p-aaa} with the difference that they have their own respective unit packet processing costs. Also, the logarithm is taken for N_{sgsn} in case of C_{p-ggsn}, N_{rnc} in case of C_{p-sgsn}, N_{bsc} in case of C_{p-rnc}, and N_{mn1} in case of C_{p-bsc} instead of N in the expression of C_{p-aaa}. C_{p-inet} represents the processing cost at the Internet and is given as:

$$C_{p-inet} = \lambda\overline{\ell}\left(\gamma_{inet} + \omega_2\left(\log_{k+1}(N_{gp}) + \frac{L}{S}\right)\right) \quad (8)$$

where γ_{inet} denotes the unit packet processing cost at the Internet, $N_{gp} = N_{ggsn} + N_{pcscf}$, and $\overline{\ell}$ is equal to 1 for IMS signaling processing cost calculation.

The processing cost between WLAN and 3G wireless cellular network for IMS data traffic C_p^{data} is:

$$\begin{aligned} C_p^{data} &= C_{p-ap} + C_{p-arg} + C_{p-wag} + C_{p-inet} \\ &\quad + 2C_{p-pcscf} + 2C_{p-scscf} + C_{p-ggsn} \\ &\quad + C_{p-sgsn} + C_{p-rnc} + C_{p-bsc} \end{aligned} \quad (9)$$

Following the same approach, we can calculate the processing cost for communication paths between WiMAX and 3G, as well as satellite and 3G for IMS traffic.

4.4 Queueing cost

For the queueing cost calculation, we first model the communication path between SN and CN as a network of M/M/1 queues [22]. The queueing cost is proportional to the total number of packets in the queueing network. The queueing cost between WLAN and 3G wireless cellular network for IMS signaling traffic C_q^{sig} is:

$$\begin{aligned} C_q^{sig} &= \omega_3(3E[n_{ap}] + 4E[n_{arg}] + E[n_{aaa}] + 3E[n_{wag}] \\ &\quad + 3E[n_{inet}] + 6E[n_{pcscf}] + 6E[n_{scscf}] \\ &\quad + 3E[n_{icscf}] + E[n_{hss}] + 3E[n_{ggsn}] \\ &\quad + 3E[n_{sgsn}] + 3E[n_{rnc}] + 3E[n_{bsc}]) \end{aligned} \quad (10)$$

where $E[n_{ap}]$, $E[n_{arg}]$, $E[n_{aaa}]$, $E[n_{wag}]$, $E[n_{inet}]$, $E[n_{pcscf}]$, $E[n_{scscf}]$ $E[n_{icscf}]$, $E[n_{hss}]$, $E[n_{ggsn}]$, $E[n_{sgsn}]$, $E[n_{rnc}]$, $E[n_{bsc}]$ denote the expected number of packets in the queue of AP, ARG, AAA, WAG, Internet, P-CSCF, S-CSCF, I-CSCF, HSS, GGSN, SGSN, RNC, and BSC, respectively. The value of $E[n_{ap}]$ is equal to:

$$E[n_{ap}] = \frac{\rho_{ap}}{1 - \rho_{ap}} \quad (11)$$

where $\rho_{ap} = \lambda_{e-ap}/\mu_{ap}$ represents the utilization at AP queue, μ_{ap} denotes the service rate at AP queue and λ_{e-ap} represents the *effective arrival rate* (in packets per second) at AP queue. That is, $\lambda_{e-ap} = \sum_{i \in N_{ap}} \lambda_i$, where N_{ap} denotes the number of active users in the AP coverage area that are engaged in communication with the AP, and hence $N_{ap} \subseteq N_{mn2}$. The effective arrival rate λ_e at a network node can be determined from the utilization at that node. Similarly, the λ_e at queues of other network nodes can be calculated and expressions can be determined for the expected number of packets at other network entities.

The queueing cost between WLAN and 3G wireless cellular network for IMS data traffic C_q^{data} is:

$$\begin{aligned} C_q^{data} &= \omega_4(E[n_{ap}] + E[n_{arg}] + E[n_{wag}] + E[n_{inet}] \\ &\quad + 2E[n_{pcscf}] + 2E[n_{scscf}] + E[n_{ggsn}] \\ &\quad + E[n_{sgsn}] + E[n_{rnc}] + E[n_{bsc}]) \end{aligned} \quad (12)$$

Following the same approach, we can calculate the queueing cost for communication paths between WiMAX and 3G, as well as satellite and 3G for IMS traffic.

5 Numerical results

In this section, we present the numerical results for the cost analysis of our proposed LCSCW2 and TCSCW2 interworking architectures. The total system signaling and data costs for IMS traffic are determined for the case when the SN is using the WLAN and the CN is in 3G wireless cellular network.

We consider a network containing two 3G BSCs, three WiMAX BSCs, 12 WLANs, and one SSB. The cell radius for 3G BSC, WiMAX BSC, and WLAN is taken to be 1,000, 700, and 50 m, respectively. Their user densities are taken to be 0.001, 0.001, and 0.008 per square meter, respectively [7, 8, 15]. The SSB is assumed to cover an area of 20 square kilometer and user density in its coverage area is taken to be 0.0005 per square meter [23]. The number of users resulting from the selection of these cell radii and user densities in different ANs are: $N_{mn1} = 5{,}000$, $N_{mn2} = 600$, $N_{mn3} = 3{,}000$, and $N_{mn4} = 10{,}000$. In our network setting, two GGSNs and two P-CSCF servers are connected to the Internet; each GGSN supports three SGSNs; each SGSN supports four RNCS, and each RNC controls five BSCs. The IP address length L and processor machine word size S are taken to be 32 bits. The system dependent constant value k is equal to 5 [21]. The wired hop distances, $d_{pcscf-ggsn}$ and $d_{sgsn-rnc}$, which involve the core 3G network entities are equal to 4, and rest of the distances are equal to 2 [7, 8, 24]. The trunked Pareto distribution is assumed for packet length with average packet length equal to 480 bytes. The inter-arrival time for packets is exponentially distributed [25].

The weighting factors, ω_1 and ω_2, corresponding to the table lookup processing cost are taken equal to 1×10^{-6} as lookup delay is increased by 100 ns for each memory access [21]. The weighting factors, ω_3 and ω_4, corresponding to queueing cost are equal to each other and are chosen such that sum of all the weighting factors is equal to 1 (i.e. $\sum_i \omega_i = 1$). We consider wireless link channels to be 9.6, 19.2 or 19.2 kbps and the wired links to be 1 Gbps. The unit transmission costs for the wired link ψ and the wireless link φ are equal to 3.84×10^{-6} and 0.1, respectively [26, 27] so that the unit transmission costs can be interpreted as typical wireless and wired link delays in seconds. The service rate μ at all the network entities is taken equal to 250 packets/sec. The unit packet processing cost for all the network entities is taken equal to 4×10^{-3} except the core 3G network entities i.e. SGSN and GGSN and the Internet for which the unit packet processing cost is taken twice as compared to other network entities in accordance with [7, 8].

For IMS data traffic, we consider audio and video sessions using different codecs which give different packet generation rates. For instance, GSM voice encoder at 13 kbps, G.726 voice encoder at 32 kbps, H.264/AVC at 56 kbps, H.264/AVC at 80 kbps, H.264/AVC at 90 kbps give packet generation rates of 4, 9, 15, and 21 packets/sec, respectively [13, 28]. The background utilization due to traffic from other sources is taken to be 0.7 for HSS and AAA server because they have to handle traffic for inter-system communications from different ANs, 0.5 for the core 3G entities i.e. SGSN and GGSN, and 0.4 for the rest of the entities. The assumption of these values of background utilizations allows us to determine λ_e at each of the network nodes.

Figure 5 shows the transmission, processing and queueing costs, as well as the total system signaling cost, for IMS signaling traffic in the LCSCW2 architecture. Results show that the ratio $C_t : C_p$ is 1 : 1.02 for signaling in the LCSCW2 architecture. With our selection of parameters, queueing cost is higher than the transmission and processing costs for signaling and lower than the transmission and processing costs for data.

Figure 6 shows the transmission, processing and queueing costs, as well as the total system signaling cost, for IMS signaling traffic in the TCSCW2 architecture. Results show that the ratio $C_t : C_p$ is 1 : 0.959 for signaling in the TCSCW2 architecture. With our selection of parameters, queueing cost is higher than the transmission and processing costs for signaling and lower than the transmission and processing costs for data.

Figure 7 shows the effect of varying IMS signaling arrival rate λ on total system signaling cost in the LCSCW2 and TCSCW2 architectures. It can be observed that the system signaling cost increases almost linearly with the increasing value of λ in both the

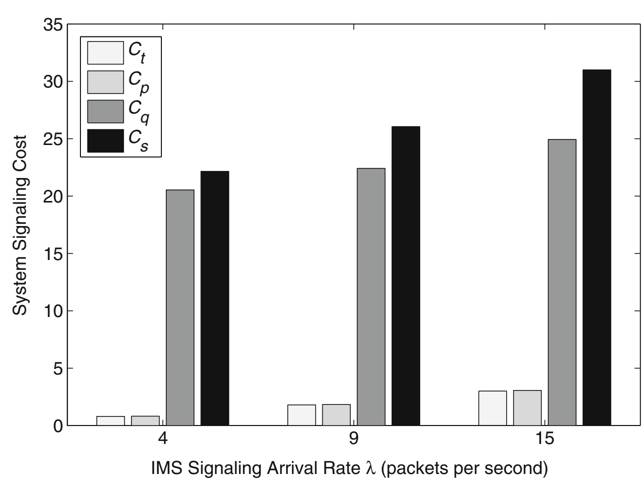

Figure 5 The breakup of system signaling cost into transmission, processing and queueing cost for different values of IMS signaling arrival rate λ in the LCSCW2 architecture

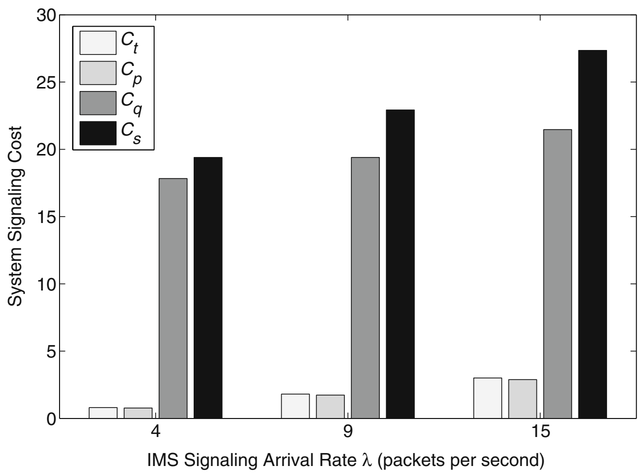

Figure 6 The breakup of system signaling cost into transmission, processing and queueing cost for different values of IMS signaling arrival rate λ in the TCSCW2 architecture

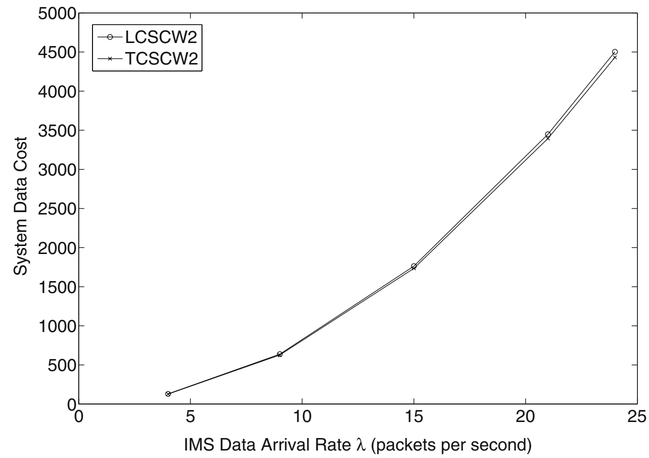

Figure 8 Effect of varying the IMS data arrival rate λ on the total system data cost between WLAN and 3G in the LCSCW2 and TCSCW2 architectures

architectures. A comparison of the system signaling cost in the LCSCW2 and TCSCW2 architectures reveals that the signaling cost in the TCSCW2 architecture is always considerably less than the LCSCW2 architecture for all values of λ. A reduction in the system signaling cost is an important achievement of the TCSCW2 architecture which justifies the deployment of PDGs. It is to be noted that signaling cost is most critical in the networks because session establishment, session resource reservation, UE registration, and vertical handoffs are achieved via the signaling. Hence, QoS can be guaranteed to an extent in the TCSCW2 architecture. In the TCSCW2 architecture, each AN has its own PDG via which the traffic is routed to IMS network without passing through the Internet and admission control mechanisms can be easily implemented to guarantee the QoS. In the LCSCW2 architecture, IMS network is reached via Internet whose background utilization can vary at different times giving an almost best effort service in the LCSCW2 architecture.

Figure 8 shows the effect of varying IMS data traffic arrival rate λ resulting from using different audio and video encoders on total system data cost in the LCSCW2 and TCSCW2 architectures. It can be observed that the system data cost increases non-linearly with the increasing value of λ in both the architectures. It can be noticed that the data cost is almost the same in both the architectures for our considered parameters. This observation implies that for data traffic both the architectures are able to provide similar kind of service. The linear increase of system signaling cost and non-

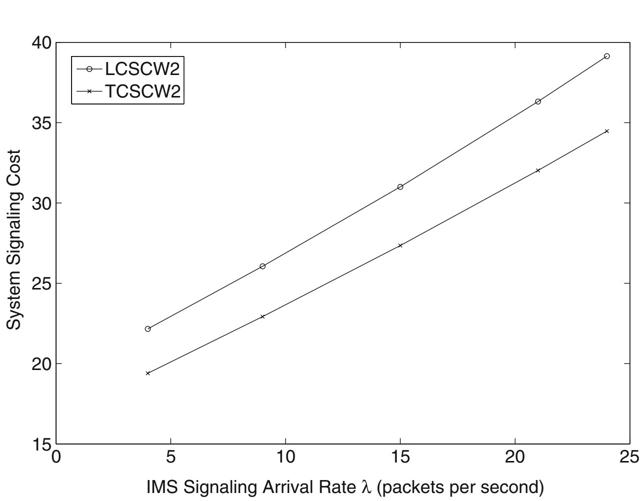

Figure 7 Effect of varying the IMS signaling arrival rate λ on the total system signaling cost between WLAN and 3G in the LCSCW2 and TCSCW2 architectures

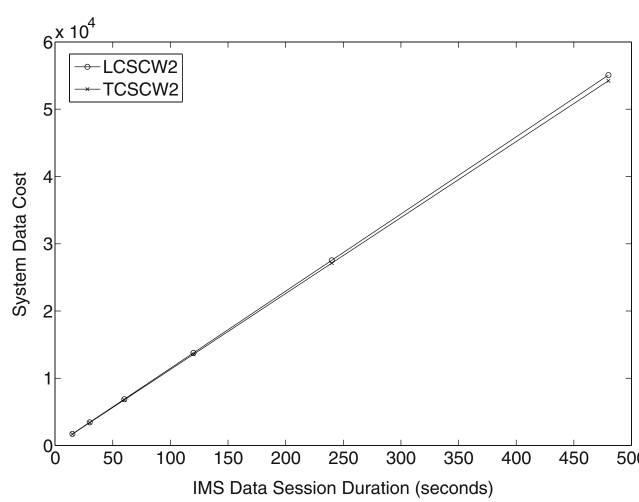

Figure 9 Effect of varying the average IMS session duration $\overline{\ell}$ on the total system data cost between WLAN and 3G in the LCSCW2 and TCSCW2 architectures

linear increase of system data cost with λ is dependent on the ratios of transmission, processing and queueing costs in the total system cost.

Figure 9 shows the effect of varying IMS session duration on total system data cost in the LCSCW2 and TCSCW2 architectures. The arrival rate λ is assumed to be to 21 packets/s. The IMS data session is run for 30, 60, 120, 240, and 480 s with the corresponding values of $\overline{\ell}$ as 315, 630, 1,260, 2,520, 5,040, and 10,080, respectively. Results show that the system data cost increases almost linearly with increasing session length for both the architectures and the system data cost is almost the same in the two architectures.

Experiments were conducted for the non-IMS traffic as well [17]. It was observed from those experiments that the TCSCW2 architecture provides significant improvement over the LCSCW2 architecture for the non-IMS traffic. Hence, the deployment of interworking functions is justified for the non-IMS traffic. For IMS session, traffic is routed to the 3G core network through IMS CSCF servers and interworking functions in the TCSCW2 architecture do not play any role in the routing of signaling and data traffic directly into the 3G core network from an AN.

6 Conclusions

In this paper, we proposed the LCSCW2 and TCSCW2 interworking architectures for 4G heterogeneous wireless networks. The LCSCW2 and TCSCW2 architectures integrates the satellite networks, 3G wireless networks, WiMAX, and WLANs. The LCSCW2 architecture supports IMS sessions, provides global coverage, and facilitates independent deployment of various access networks. We also proposed a cost model to determine the associate cost for the IMS signaling and data traffic in the LCSCW2 architecture. We presented the numerical results for the system cost, as well as the transmission, processing, and queueing costs under different arrival rates and session lengths for LCSCW2 and TCSCW2 architectures.

For future work, one direction is to consider a hybrid approach which combines the benefits of LCSCW2 and TCSCW2 architectures. The hybrid architecture is envisioned to provide maximum reliability and QoS support. The hybrid architecture should be capable of operating in two modes i.e. the tightly coupled mode and the loosely coupled mode. In this architecture, there will always be two paths available for traffic, one arising from LCSCW2 and the other from TCSCW2. The network needs to choose a path from the two available ones that promises to optimize system performance by providing the least inter-system communication cost.

Acknowledgements This work was supported by Bell Canada and Natural Sciences and Engineering Research Council of Canada (NSERC).

References

1. 3GPP (2006) IP multimedia subsystem (IMS); Stage 2 TS 23.228 (v7.3.0), March
2. Geer D (2005) Building converged networks with IMS technology. IEEE Comput 38(11):14–16, November
3. Buddhikot M, Chandranmenon G, Han S, Lee YW, Miller S, Salgarelli L (2003) Integration of 802.11 and third-generation wireless data networks. In: Proc. of IEEE conference on computer communications (INFOCOM), San Francisco, CA, USA, March–April
4. Salkintzis AK, Fors C, Pazhyannur R (2002) WLAN-GPRS integration for next-generation mobile data networks. IEEE Wirel Commun 9(5):112–124, October
5. Ruggeri G, Iera A, Polito S (2005) 802.11-based wireless LAN and UMTS interworking: requirements, proposed solutions and open issues. Elsevier Comput Netw 47(2):151–166, February
6. Song J-Y, Lee HJ, Lee S-H, Lee S-W, Cho D-H (2006) Hybrid coupling scheme for UMTS and wireless LAN interworking. Int J Electron Commun 4:2247–2251, July
7. Liu C, Zhou C (2005) An improved interworking architecture for UMTS-WLAN tight coupling. In: Proc. of IEEE Wireless Communications and Networking Conference (WCNC), Atlanta, GA, USA, March
8. Liu C, Zhou C (2005) HCRAS: a novel hybrid internetworking architecture between WLAN and UMTS cellular networks. In: Proc. of IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, USA, January
9. Wong KD, Varma VK (2003) Supporting real-time IP multimedia services in UMTS. IEEE Commun Mag 41(11): 148–155, November
10. Mahmood HS, Gage B (2003) An architecture for integrating CDMA2000 and 802.11 WLAN networks. In: Proc. of IEEE vehicular technology conference (VTC), Orlando, FL, USA, October
11. Evans BG, Werner M, Lutz E, Bousquet M, Corazza GE, Maral G, Rumeau R, Ferro E (2005) Integration of satellite and terrestrial systems in future multimedia communications. IEEE Wireless Commun 12(5):72–80, October
12. Nguyen-Vuong QT, Fiat L, Agoulmine N (2006) An architecture for UMTS-WIMAX interworking. In: Proc. of 1st international workshop on broadband convergence networks (BcN'06), Vancouver, British Columbia, Canada, April
13. Kim D, Ganz A (2005) Architecture for 3G and 802.16 wireless networks integration with QoS support. In: Proc. of 2nd international conference on quality of service in heterogeneous wired/wireless networks (QShine'05), Orlando, FL, USA, August
14. Fan L, Sheriff R, Gardiner J (2000) Satellite-UMTS service provision using IP-based technology. In: Proc. of IEEE vehicular technology conference (VTC), Tokyo, Japan, May

15. Mohanty S, Xie J (2007) Performance analysis of a novel architecture to integrate heterogeneous wireless systems. Comput Netw 51(4):1095–1105, March
16. 3GPP (2006) 3GPP system to wireless local area network (WLAN) interworking; system description. TS 23.234 (v7.2.0), June
17. Munir A (2007) Analysis of interworking architectures for IP multimedia subsystem. Master's thesis, University of British Columbia, Vancouver, BC, Canada, April
18. 3GPP (2006) Signalling flows for the IP multimedia call control based on session initiation protocol (SIP) and session description protocol (SDP); stage 3 (Release5). TS 24.228 (v5.15.0), September
19. Camarillo G, Garcia-Martin M-A (2004) The 3G IP multimedia subsystem (IMS): merging the Internet and the Cellular Worlds. John Wiley and Sons
20. Calhoun P, Loughney J, Guttman E, Zorn G, Arkko J (2003) Diameter base protocol. RFC 3588, September
21. Lampson B, Srinivasan V, Varghese G (1999) IP lookups using multiway and multicolumn search. IEEE/ACM Trans Netw 7(3):324–334, June
22. Medhi J (2003) Stochastic models in queueing theory. Academic Press, An imprint of Elsevier Science
23. Ayyagari D, Ephremides A (1998) A satellite-augmented cellular network concept. ACM Wirel Netw 4(2):189–198, February
24. Pack S, Choi Y (2004) A study on performance of hierarchical mobile IPv6 in IP-based cellular networks. IEICE Trans Commun E87-B(3):462–469, March
25. Tarchi D, Fantacci R, Bardazzi M (2006) Quality of service management in IEEE 802.16 wireless metropolitan area networks. In: Proc. of IEEE international conference on communications (ICC), Istanbul, Turkey, June
26. Das SK, Lee E, Basu K, Sen SK (2003) Performance optimization of VoIP calls over wireless links using H.323 protocol. IEEE Trans Comput 52(6):742–752, June
27. Kurose JF, Ross KW (2003) Computer networking, a top-down approach featuring the internet. Addison-Wesley
28. Fröjdh P, Horn U, Kampmann M, Nohlgren A, Westerlund M (2006) Adaptive streaming within the 3GPP packet-switched streaming service. IEEE Netw 20(2):34–40, March–April

Arslan Munir received the B.Sc. degree from the University of Engineering and Technology (UET), Lahore, Pakistan, in 2004, and the M.A.Sc. degree from the University of British Columbia, Vancouver, BC, Canada, in 2007. He is currently a software development engineer at Mentor Graphics. His research interests are in design and performance evaluation of heterogeneous wireless networks.

Vincent W. S. Wong received the B.Sc. degree from the University of Manitoba, Winnipeg, MB, Canada, in 1994, the M.A.Sc. degree from the University of Waterloo, Waterloo, ON, Canada, in 1996, and the Ph.D. degree from the University of British Columbia (UBC), Vancouver, BC, Canada, in 2000. From 2000 to 2001, he worked as a systems engineer at PMC-Sierra Inc. He joined the Department of Electrical and Computer Engineering at UBC in 2002 and is currently an Associate Professor. His research interests are in resource and mobility management for wireless mesh networks, wireless sensor networks, and heterogeneous wireless networks. Dr. Wong is an associate editor of the IEEE Transactions on Vehicular Technology and an editor of KICS/IEEE Journal of Communications and Networks. He serves as TPC member in various conferences, including IEEE Globecom'08, ICC'09, and Infocom'09. He is a senior member of the IEEE and a member of the ACM.