Reliable Integrated Architecture for Heterogeneous Mobile and Wireless Networks

Christian Makaya and Samuel Pierre
Mobile Computing and Networking Research Laboratory (LARIM)
Department of Computer Engineering, Ecole Polytechnique de Montreal
P.O. Box 6079, Station Centre-ville, Montreal, Quebec, H3C 3A7, Canada
Email: {christian.makaya, samuel.pierre}@polymtl.ca

Abstract— **The major trend in next-generation or 4G wireless networks (NGWN/4G) is the coexistence of diverse but complementary architectures and wireless access technologies. In this context, an appropriate integration and interworking of existing wireless networks are crucial to allow seamless roaming across those networks. Several integrated architectures have been proposed for 3G cellular networks and wireless local area networks (WLANs) by both third generation wireless initiatives, 3GPP and 3GPP2. However, the proposed architectures have several drawbacks, the most significant being the absence of quality of service (QoS) guarantees, seamless roaming and service continuity. This paper proposes a novel architecture, called *Integrated InterSystem Architecture* (IISA), which enables the integration and interworking of various wireless networks and hide their heterogeneities from one another. The IISA architecture aims provisioning of guaranteed seamless roaming and service continuity across different access networks. Performance evaluation shows that IISA together with the proposed handoff management scheme provide significant gains than existing interworking architectures and mobility management protocols.**

Index Terms— **Interworking architecture, seamless roaming, mobility management, quality of service.**

I. Introduction

Next-generation or 4G wireless networks (NGWN/4G) are expected to exhibit heterogeneity in terms of wireless access technologies and services. The advantages of 3G cellular networks such as UMTS and 1xEV-DO/DV, consist of their global coverage while their weaknesses lie in their bandwidth capacity and operational costs. On the other hand, WLAN technology such as IEEE 802.11 offers higher bandwidth with low operational costs, although it covers relatively short range. Moreover, technological advances in evolution of portable devices have made possible the support of different radio access technologies (RATs). This has raised much interest in the integration and interworking of 3G wireless networks and WLAN, in order to benefit of their respective potentials.

The integration of these systems seems unavoidable due to potential benefits of their complementarity and will be the path toward the design of NGWN/4G instead of putting efforts into developing new radio interfaces and technologies [1]. The purpose of the integration of different networks is to unify the advantages of these systems and at the same time to minimize the disadvantages. This allows a great market opportunity. Conceptually, NGWN/4G architecture can be viewed as many overlapping wireless access domains, as shown in Fig. 1. Furthermore, heterogeneity in terms of RATs and network protocols in NGWN/4G asks for common interconnection element. Since the IP (Internet Protocol) technology enables the support of applications in a cost-effective and scalable way, it is expected to become the core or backbone network of NGWN/4G [2]. Thus, current trends in communication networks evolution are directed towards the *all-IP* concept, in order to hide heterogeneities and to achieve convergence of different access networks.

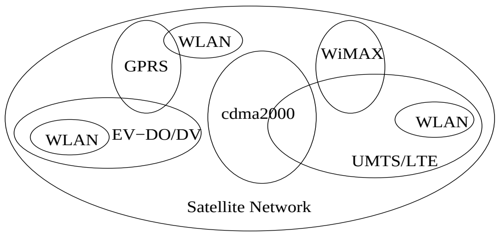

Figure 1. Overview of 4G/NGWN network architecture.

The integration of 3G cellular networks and WLAN may be done at several points. Two major architectures (*loose* and *tight coupling*) for 3G/WLAN interworking have been proposed by both 3G wireless network initiatives, 3GPP and 3GPP2, for their respective system [3], [4]. However, this integration brings new challenges such as selection of integration point, mobility management, interworking, QoS guarantees and security issues. These challenges are key issues in order to support global roaming and service continuity of mobile nodes (MNs) across various networks in an efficient way.

This paper proposes a novel architecture, called *Integrated InterSystem Architecture* (IISA), based on 3GPP/3GPP2-WLAN interworking models, to integrate the existing wireless systems and hide their heterogeneities from one another. The main purpose of the

This paper is based on "An Interworking Architecture for Heterogeneous IP Wireless Networks" by C. Makaya and S. Pierre, which appeared in the Proceedings of the 3rd International Conference on Wireless and Mobile Communications (ICWMC), Guadeloupe, French Caribbean, March 2007. © 2007 IEEE.
This work was supported in part by the NSERC/Ericsson Industrial Research Chair under grant IRC252720-01.

IISA is to enable QoS guarantees, seamless roaming and service continuity for real-time applications in heterogeneous IPv6-based wireless environments. The remainder of this paper is organized as follows. Section II offers an overview of the basic concepts and inherent challenges of designing an integrated architecture are described. After that, related work on the interworking and integration in heterogeneous wireless networks are presented in Section III. The proposed interworking architecture, IISA, is presented in Section IV, followed by performance evaluation in Section V before to conclude the paper.

II. Basic Concepts Overview

According to heterogeneity in NGWN, provisioning of seamless mobility and service continuity (i.e., minimum service disruption during roaming) based on intelligent and efficient mechanisms is crucial. However, maintaining uninterrupted session while the physical interface is changing constitutes a complex task in NGWN/4G. Handoff management is defined as a capability for managing the mobility for a mobile node (MN) in active state. An evident way to achieve roaming across networks of different service providers or operators is by using service level agreements (SLAs). However, this approach is not always feasible. In fact, the increasing number of wireless networks and service providers make it impractical for an operator to have direct SLAs with all other operators. Moreover, network operators are reticent to make their databases available to other operators.

A reliable integrated and interworking architecture for NGWN/4G should handle specific requirements and have the following main features [2]:

- *economical*: to ensure economical and rapid deployment, the architecture should use the existing infrastructures as much as possible and minimize the usage of new infrastructures;
- *scalable*: integration of any number of wireless systems of both existing and future service providers should be supported by the architecture and be able to provide fault tolerance;
- *seamless mobility*: to eliminate connection interruptions and the QoS degradation during intersystem or intrasystem roaming, the architecture should support seamless mobility or roaming;
- *security*: the architecture should provide a level of security and privacy which is equivalent or better than the existing wireless and wired networks.

The above challenges and requirements show that it is very hard to have a single integrated architecture which is appropriate for all interworking scenarios and satisfy both network operators and wireless internet service providers. It is very difficult to forecast which interworking architecture will dominate in the market, since selection of model is not based only on performance criterion, but on its cost and its respective profits. A practical solution can be achieved by a certain tradeoff of the above requirements.

Six 3G/WLAN interworking scenarios and their requirements have been defined in [3] and [4], in order to provide a proper background for interworking architecture design. With the particular characteristics of WLAN and 3G wireless networks, two scenarios present significant technical challenges: services continuity and seamless roaming provisioning.

III. Related Work

Several 3G wireless networks and WLAN interworking architectures are available in the literature. Two majors of them have been proposed by 3GPP and 3GPP2, called *loose coupling* and *tight coupling* [3]. With the tight coupling approach, WLAN appears to the 3G wireless core network as one of the 3G wireless radio access network (RAN). MNs must implement both 3G networks and WLAN interfaces at lower layer of TCP/IP protocols stack. The 3G wireless networks protocol stack should be implemented on top of WLAN technology in MNs' devices. Although, the tight coupling allows easy control of QoS for time-sensitive application, it leads to several drawbacks such as high cost and complexity. Both technologies should be owned by the same wireless operators, MNs' devices and configurations should be modified. Moreover, with tight coupling, traffic from WLAN flows into 3G wireless core network and leads to capacity problems. In fact, 3G wireless core network nodes cannot accommodate the bulk data traffic from WLAN.

On the other hand, with loose coupling, different networks are deployed independently and the data paths are completely separated between WLAN and 3G networks. Hence, the loose coupling enables several advantages in terms of low cost and less complexity: independent traffic engineering, deployment and ownership of both technologies, fewer networks and mobile devices modifications, etc. However, the loose coupling may not guarantee service continuity to other access networks during handoff, because it has higher handoff latency and packet loss. In fact, the QoS provisioning with loose coupling depends on the Internet QoS status. The hybrid coupling is also proposed in the literature and differentiates the data path according to the type of traffic [5], [6]. With the hybrid coupling, the real-time traffic uses the path based on the tight coupling while non-real time traffic uses the path based on the loose coupling. By combining advantages of 3G wireless networks and WLAN, the hybrid coupling can provide seamless handoff in terms of low packet loss and low delay. However, some drawbacks of the tight and loose coupling still exist in the hybrid coupling.

An architecture for the next-generation all-IP-based wireless systems is proposed in [2] and called Architecture for Ubiquitous Mobile Communications (AMC). Two new entities, Network Interworking Agent (NIA) and Interworking Gateway (IG) are introduced. The QoS guarantee is not taken into account in AMC and deployment of NIA and IG require extra cost. Moreover, the AMC architecture does not provide appropriate handoff decision mechanism to take into account heterogeneity of access networks. Other works have been done for the interworking of heterogeneous 3G cellular networks [7],

[8] but not for IP-based wireless networks or different types of access networks technologies. These integration schemes are based on the deployment of a gateway, which takes care of interworking issues, between each pair of networks. Adding a gateway at each boundary between two systems will increase deployment costs and arises scalability issues as well as transparency amongst heterogeneous access technologies.

The choice of an optimal interworking architecture is determined by some factors. For example, if the wireless network is composed by a large number of WLAN and 3G wireless operators, the loosely coupled architecture would be the best choice. On the other hand, if the WLAN network is owned exclusively and operated by a 3G wireless operator, the tightly or hybrid coupled architecture might become a more attractive option. The loose coupling approach offers more advantages than tight coupling, with virtually no drawbacks and it is the most advocated interworking scheme in the literature [9]. Although most proposed schemes offer several advantages, they continue to be hindered by certain drawbacks, the most significant being the absence of the QoS guarantee and seamless roaming support.

IV. PROPOSED INTEGRATION FRAMEWORK

In order to guarantee an ubiquity or *always best connected* [10] features to mobile users, this paper proposes an interworking architecture, called *Integrated InterSystem Architecture* (IISA). Its key objective is to allow seamless service continuity across various RATs. Instead of developing new infrastructures, IISA extends existing infrastructure to tackle the integration and interworking issues. For the sake of simplicity, only UMTS, CDMA2000 and WLAN networks are illustrated in the IISA architecture in Fig. 2(a). However, IISA may integrate any number of RATs and mobile devices may be equipped with any number of interfaces.

A. Integrated InterSystem Architecture

The proposed architecture is shown in Fig. 2(a) and is based on adaptive loose coupling model. With the IISA architecture, various integrated networks appear as peer-networks. The IISA uses hierarchical architecture and is IPv6-based, i.e., it implements Mobile IPv6 (MIPv6) [11] and Hierarchical MIPv6 (HMIPv6) [12] functionalities. MIPv6 has been proposed for mobility management at the IP layer. However, MIPv6 has some well known drawbacks such as signaling traffic overhead, higher handoff latency and packet loss rate thereby causing user-perceptible deterioration of real-time traffic. In order to address these problems, HMIPv6 was proposed to handle handoff locally through a special node called Mobility Anchor Point (MAP). Hence, the amount of MIPv6 signaling outside the MAP domain is limited and the location update delay is reduced.

In the IISA architecture, a novel entity, *Interworking Decision Engine* (IDE), is introduced to enable interworking between different networks. The IDE may be under the responsibility of the third-party service provider (owned by one or multiple operators with SLAs among them) like it is the case for GPRS Roaming eXchange (GRX) in GPRS networks [13]. Then, the network operator needs to establish only one direct SLA with the IDE manager instead of establishing individual SLAs with all other operators. Usage of the IDE could be seen as a value-added services that operators offer to their subscribers to allow global roaming. If necessary, an IDE operator will be responsible for making additional agreements with other IDE managers.

To provide the support of IPv6-based mobility management protocols, some functional entities of UMTS/CDMA2000 networks are extended. Hence, the Serving GPRS (General Packet Radio Service) Support Node (SGSN) and Packet Control Function (PCF) are extended with the functionalities of an access router (AR) and are called *Access Edge Node* (AEN). Similarly, the Gateway GPRS Support Node (GGSN) and Packet Data Serving Node (PDSN) are extended with MAP (Mobility Anchor Point) and interworking functionalities (to enable message formats conversion, QoS requirements mapping, etc.) and are called *Border Edge Node* (BEN). The BEN has the information for ARs such as IP address, subnet prefix, link address within its domain. The WLAN interworking gateway (WIG) acts as a route policy element, ensuring message format conversion. Extended functionalities can be integrated into existing network entities or implemented separately. We advocate for the first scenario since it is easily deployed and managed.

The IISA adds the AAA (authentication, authorization and accounting) linkage and supports of IPv6-based mobility management scheme when it is not available. Interworking of different access networks is required for an efficient integration. The mapping between home location register or home subscriber server (HLR/HSS) in 3G wireless networks and AAA server in WLAN is required to allow execution of authentication and billing when user roams across both technologies. In the IISA architecture, authentication is done by combining AAA protocol and context transfer or token-based approach. We make distinction between home AAA server (AAAH) located in the MN's home network and local AAA server (AAAL) located in the foreign network. IISA allows the separation between the control plane (signaling traffic) and the transport plane (data traffic). In fact, only signaling traffic goes through the IDE, but not data packets. This avoid the IDE to become a potential bottlenecks or point of failure.

B. Interworking Decision Engine

The Interworking Decision Engine (IDE) is introduced to manage handoff regardless of wireless access technologies, operators and wireless service providers. It is designed for the purpose of exchanging required information between heterogeneous wireless systems in order to reduce signaling traffic and services disruption during handoff. Specifically, the IDE handles AAA and mobility

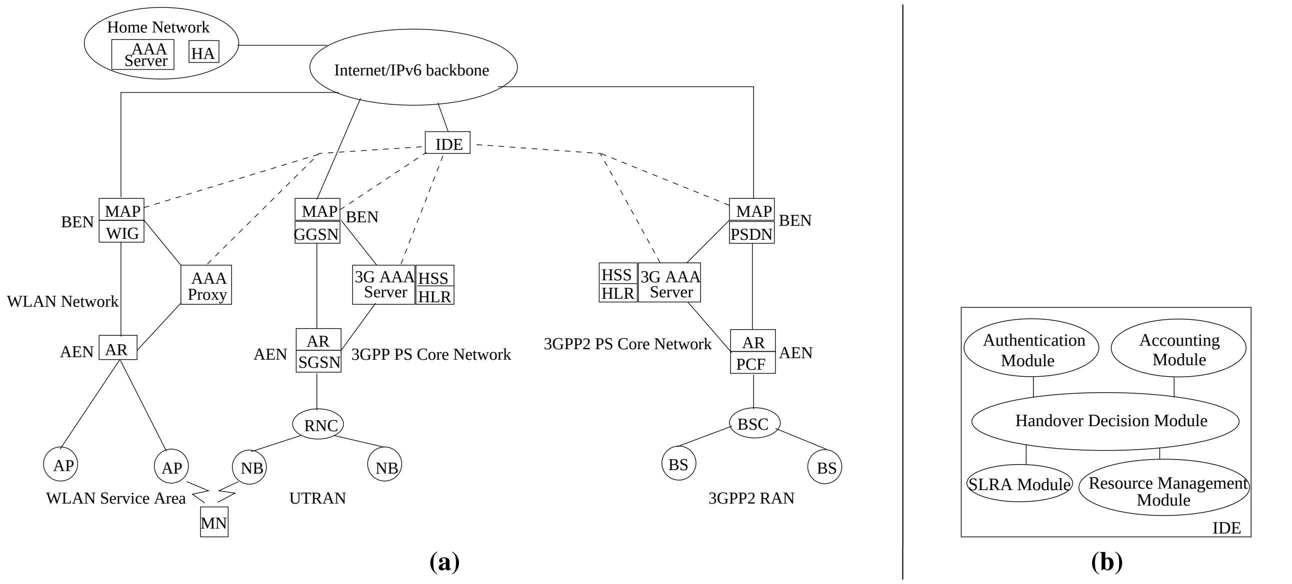

Figure 2. The proposed networks integration framework: (a) Integrated InterSystem Architecture (IISA); (b) Interworking Decision Engine (IDE).

management for intersystem and/or inter-domain roaming. The IDE makes policy decisions and provides mediation between different service or network providers. To reduce IDE's load, the IDE is involved only for intersystem and/or inter-domain handoff process and it manages only control signaling traffic; users' data packet traffic does not go through the IDE. In fact, the IDE is in a control plane while the MAP/BEN handles the actual traffic, thus it is in the transport plane. By separating the control and transport plane, the IISA architecture becomes flexible for adding new services, and offers easy interworking with legacy networks.

Furthermore, to enable the scalability of the IISA architecture, if the number of mobile users that requires intersystem and/or inter-domain handoff increases or if the number of heterogeneous wireless systems increases, the IDE can be deployed in hierarchical or distributed framework. For roaming user with ongoing session, the IDE allows the reduction of association and authentication delays. Since the IDE is closer to foreign network than home network to foreign network; then, handoff process execution is speed up. To allow easy deployment of the IDE, it may be placed at a control point within a signaling network or core network, for example in the Internet.

Logical components of the IDE are illustrated in Fig. 2(b). The *Authentication Module* (AuM) is used to authenticate users moving across different wireless networks and it avoids the need for direct security agreements or associations between foreign and home network. When an MN enters into a new domain for the first time, authentication and authorization procedures are performed between the foreign and home network through the IDE. The credentials information are then stocked in the IDE and a token is provided to MN for further authentication and authorization needs. Thus, when an MN moves to another foreign network, an end-to-end re-authentication is not necessary. The MN will just use its token and send it to the IDE for validation. The AuM emulates HSS/HLR functionalities for WLAN's subscribers to enable usage of 3G wireless network legacy authentication and location update procedures when WLAN subscriber roams into 3G wireless network. The AuM maintains an entry list until lifetime is expired. If the lifetime expires, the entry is removed. However, the lifetime may be refreshed by a request to AAAH server. The WLAN AAA server/proxy routes the AAA messages to appropriate 3G AAA server through AuM/IDE and vice versa.

The *Accounting Module* (AcM) enables billing between different wireless networks. It acts as common billing/charging system between various network operators. The AcM collects accounting information received from AAA server of the foreign network per-user based on the charging policy of the foreign network operator. It converts if necessary call detail records of the foreign network before to forward this information to AAAH server for billing purposes. Charging information associated to resource usage is stored in AcM. The CIBER (Cellular Intercarrier Billing Exchange Roamer Record) protocol may be used for the exchange of roaming billing information among wireless operators through the IDE.

Usually, different administrative domains have different QoS policies for resources allocation. Then, when an MN moves between two different administrative domains, the QoS re-negotiation may be required. This re-negotiation will be based on SLAs between both domains through the IDE. Hence, the *Resource Management Module* (RmM) enables QoS mapping, fast transfer of user profile and QoS parameters between different domains during handoff. The RmM handles also operations for bandwidth management that offers resource allocation, policy enforcement and call admission control. The QoS mapping and the mechanism by which the IDE allocates resources to an MN, and decides to admit a new request is outside the scope of this paper. However, we assume that the IDE is endowed with intelligence and can perform the following operations: translation of signaling message formats between different networks, conversion of higher transmission rate to lower rate, translation of QoS parameters and information, etc.

The *SLRA Module* stores information about service providers or network operators who have SLAs and roaming agreements (RAs) with the IDE manager. The *Handover Decision Module* (HdM) is used when inter-system and/or inter-domain handoff should be granted or not. In other words, it enables support of roaming and handoff for mobile users. The HdM module verifies with SLRA module the existence of agreements with MN's home network. Moreover, the HdM decides the best available network in case of network-controlled handoff and enables efficient load balancing. The HdM includes also the MAP functionalities for mobility management of users who perform inter-domain/system handoff. If some lawful operations such as legal intercept are required, a decision module to handle them may be included as MAP functionalities of the HdM.

C. Registration and Roaming Procedures

To avoid the additional signaling overhead due to the execution of AAA procedure each time an MN performs handoff and request registration, a token-based approach is proposed. During roaming within the MAP/BEN domain of access networks having agreements with the IDE, an MN presents a token, which it obtains from the IDE after its first successful registration in the foreign network, to the MAP/BEN or AR/AEN. The token includes security association parameters for secure tunnel sets up and context transfer. This yield a lower registration latency than performing authentication and authorization check with the AAAH server. If the MAP/BEN or AR/AEN verifies the token successfully, it initiates the authorization process. The home agent (HA) functionalities related to MN authentication, distributing keying materials, security association, context transfer and mobility management are delegated to the IDE during MN roaming. Subsequent movements are handled either by the MAP/BEN and AAAL server or by the IDE whether movement is intrasystem or intersystem.

When an MN detects it is moving out of its residence area, for example from L2 trigger [14], the MN selects the best target network from relevant information received through network entities. The handoff decision function proposed in [15] may be used since it is more appropriate in heterogeneous networks environments. After the subnet selection, the MN initiates authentication and authorization procedures which are combined with MIPv6 registration procedure. The request sent by an MN to the MAP/BEN, allows the latter to know that the MN is a roaming user. The MAP/BEN can then start handoff procedure execution by determining if an intersystem or inter-domain is required for the MN. In case of inter-system or inter-domain handoff, the MAP/BEN forwards registration request to the IDE. The latter determines if the MN may be granted the permission to access the foreign network according to SLAs its home network operator has established with the IDE manager. If SLAs exist, the IDE performs registration request along to authentication and authorization. After a successful registration, the MN can start communication through the new subnet. The handoff procedure is illustrated in Fig. 3.

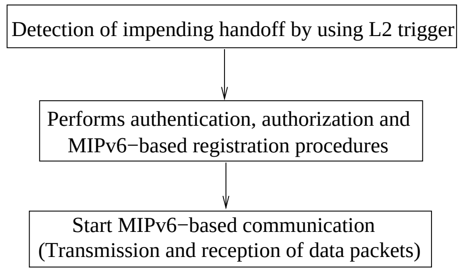

Figure 3. Handoff execution over IISA architecture.

D. Context Transfer and Binding Update

To achieve seamless mobility across various access technologies and networks, an MN needs information about the wireless network to which it could attach. Also, it is necessary to transfer information (context transfer) related to an MN from the current access router to the next one. To enable these procedures, the Candidate Access Router Discovery protocol (CARD) [16] and the Context Transfer Protocol (CXTP) [17] have been proposed. Their key objectives consist of reducing latency, packet losses, and avoiding the re-initiation of signaling to/from an MN from the beginning during an handoff.

However, context transfer is not always possible. For example, when an MN moves across different administrative domains, the new network may require the MN to re-authenticate and perform signaling from the beginning rather than accepting the transferred context. Also, MNs must periodically monitor the received signal strength from neighboring AP/ARs and construct neighbor network information table. Moreover, entities exchanging contexts must authenticate each other, which could turn into a tedious procedure in 4G/NGWN. We propose an adaptive context transfer by using the IDE as a network anchor entity.

When the MAP/BEN receives a handoff request message, it transmits a Context Transfer Data (CTD) message, to the NAR/AEN containing feature contexts. Example of features contained in CTD message are QoS context information, header compression, security and AAA parameters. This paper mainly focuses on QoS context information. The routers extract this QoS context information, and according to context received, the intermediate router reserves corresponding resources and updates the path information. If the MAP/BEN has no context pertaining to the concerned MN, the new MAP/BEN sends a Context Transfer Request (CTReq) message to the IDE in order to obtain session management parameters for this MN and to establish traffic bearers on the new path. In response to a CTReq message, the IDE transmits a CTD message that includes the MN's previous IP address and feature contexts. When the MAP/BEN receives a CTD message, it installs the contexts as received from the IDE. The MAP/BEN includes the CTD message within the handoff request message and forwards it to the new AR (NAR) or AEN.

When the NAR/AEN receives the CTD message, it can generate a CTD Reply (CTDR) message optionally to report the processing status of the received contexts and piggybacks this message into the handoff reply message. The binding update (BU) procedure is performed by the NAR/AEN on the behalf of MNs. In fact, an AR/AEN acts as a proxy: copies a BU list of an MN in its cache and manages this list (e.g., lifetime entries) in the same way as the original is managed by the MN. The AR/AEN cache copy must be updated periodically according to the original BU list of the MN. The BU list contains information about used home address and care-of address (CoA), IPv6 address of CNs, sequence number, lifetimes, and state of retransmissions. When the BU list lifetimes cached in AR/AEN is about to expire, the AR/AEN may send a BU list renewal request to the MN. The BU list renewal is performed in the same way as a classical BU refresh [11]. Then a separate out-of-band messages from MN to NAR/AEN are avoided, thus reducing signaling traffic overhead.

V. Performance Evaluation

We analyze the handoff procedure by considering authentication and binding update delay in order to show the effectiveness of our proposal over existing schemes. In other words, we define total cost (C_T) as the sum of the authentication cost (C_A) and the registration and roaming cost (C_R): $C_T = C_A + C_R$. Note that, the authentication cost is the same with traditional approaches at the first movement to the new MAP/BEN domain. However, the rest of processing is more effective since the authentication is performed by the IDE. Let T_{AC} be the sum of the L2 (link layer switching) handoff latency, router discovery delay and duplicate address detection (DAD) delay; $T_{X,Y}$ be the one-way transmission delay between nodes X and Y. We compare cases when IPv6-based mobility schemes such as MIPv6 and HMIPv6 are used in traditional interworking architectures with our handoff management protocol in the IISA architecture.

A. Handoff Latency and Packet Loss Rate

With MIPv6, whenever the MN across subnet boundary, it must register and authenticate with the HA/AAAH in home network first. After that, the return routability procedure [11] is performed with all active correspondent nodes (CNs) followed eventually by binding update to CNs. Hence, the handoff latency for any movement, when MIPv6 is used, is given as follows:

$$D_{MIPv6} = T_{AC} + 2T_{MN,HA} + T_{RR} + 2T_{MN,CN} \tag{1}$$

where $T_{RR} = 2\max[(T_{MN,HA} + T_{HA,CN}), T_{MN,CN}]$ is the delay of return routability procedure. We assume that processing delay and routing table lookup delay are negligible compared to access and to transmission delay.

An MN performs two types of binding update with HMIPv6: local and global. Global binding update occurs when an MN moves out of its MAP domain while local binding update is performed when an MN changes its current IP address within a MAP domain. Hence, for global binding update, the MN first registers with a local MAP and thereby obtains a regional care-of address (RCoA) on the MAP's link, then registers this RCoA to HA and CNs. Hence, the handoff latency for local and global binding update, when HMIPv6 is used, is given as follows:

$$\begin{aligned} D^{l}_{HMIPv6} &= T_{AC} + 2T_{MN,MAP} \\ D^{g}_{HMIPv6} &= 2T_{MN,MAP} + D_{MIPv6}. \end{aligned} \tag{2}$$

The proposed roaming management scheme in this paper is based on HMIPv6. Hence, we obtain same handoff latency. However, the main difference is with authentication process. In fact authentication procedure for MIPv6 and HMIPv6 is similar while the authentication in IISA is delegated to the IDE. The associated authentication delays are given as follows:

$$\begin{aligned} D^{A}_{MIPv6} &= D^{A}_{HMIPv6} = 2T_{MN,HA} \\ D^{A}_{IISA} &= 2T_{MN,IDE}. \end{aligned} \tag{3}$$

With IPv6-based mobility management protocols, packet loss occurs during handoff or service disruption latency. In fact, the number of packet loss is proportional to handoff latency.

B. Processing Load of the IDE

Wireless overlay networks are subdivided into low-tier (e.g., WLAN) and high-tier (e.g., 3G wireless network) [18]. Roaming between low-tier and high-tier networks refers to vertical or intersystem handoff. To analyze the load incurred at the IDE, we assume that high-tier networks overlap with the low-tier networks and users are uniformly distributed. The centralized nature of the IDE may arise scalability issue of IISA. However, by separating signaling and data traffic, the IDE is able to support a huge number of users. Furthermore, centralized controllers have been successfully employed in the literature and it is certainly scalable. Otherwise, as stated above, if needed, the IDE may be deployed hierarchically or in a distributed framework.

Recall that with MIPv6, each subnet crossing results in a binding update to the HA. Moreover, during refresh time period, each MN sends out a refresh request to the HA. Let N_h and N_l be the number of high-tier and low-tier networks in the service or coverage area (e.g., one city), respectively. The user density is denoted by ρ_h and ρ_l in high-tier and low-tier networks. Thus, the processing load at the HA with MIPv6 scheme is given by (4).

Let N_s be the total number of subnets in a high-tier network, $N_h \leq N_s$, ν_l (resp. ν_h) stands for the proportion of subscribers in low-tier (resp. high-tier) network away from their home network, P_{BU} the processing time for

$$L_{HA} = P_{BU}\frac{[N_l \rho_l v_l L_l + N_s \rho_h v_h L_s]}{\pi} + P_{BR}\frac{[\nu_l \rho_l A_l N_l + \nu_h \rho_h A_h N_h]}{T_{HA}} \quad (4)$$

$$L_{IDE} = P_{PU}\frac{[N_l \rho_l v_l L_l + N_h \rho_h v_h L_s]}{\pi} + P_{PR}\frac{\left\lceil \frac{\nu_l \rho_l A_l N_l}{\varepsilon_l} \right\rceil + \left\lceil \frac{\nu_h \rho_h A_h N_h}{\varepsilon_h} \right\rceil}{T_{IDE}} \quad (5)$$

registration update message and P_{BR} the processing time for binding refresh message. T_{HA} and T_{IDE} denote the binding lifetime period at the HA and the IDE, while A_l and A_h indicate the coverage area of low-tier and high-tier networks. On the other hand, v_l and v_h are the average speed of an MN in low-tier and high-tier networks, while L_l is the perimeter of low-tier network and L_s the perimeter of a subnet in high-tier network.

In IISA, binding refresh and binding update are performed locally at the MAP/BEN and not to the IDE as long as an MN moves within the MAP/BEN domain or performs intrasystem handoffs. However, during the refresh time period T_{IDE} the MAP/BEN sends one Request or Reply message to the IDE for a given number of MNs. We denote ε_l the number of these MNs for low-tier networks and ε_h for high-tier networks. Therefore, when intersystem and/or inter-domain handoff occurs, path updates are required. Thus, the IDE processing load is given by (5). P_{PU} stands for the processing time for path updates and P_{PR} is the processing time for path refresh message.

C. Simulation Results

For performance analysis, we consider random-walk mobility model and the following system parameters: $T_{MN,HA} = 30$, $T_{MN,CN} = 20$, $T_{MN,IDE} = 8$, $T_{CN,HA} = 10$, $T_{AC} = 100$ and $T_{MN,MAP} = 6$. Let μ be the subnet crossing rate; we assume that average subnet residence time is $1/\mu = 10$ seconds and boundary crossing probability $p = 0.65$, when they are not considered as variable parameters. Parameters values used in performance evaluation are given in Table I.

The network topology considered for analysis is illustrated in Fig. 4. We assume the distance (number of hops) between different domains to be equal, i.e., $c = d = e = f = 10$ and set $a = 1$, $b = 2$, and $g = 4$. All links are supposed to be full-duplex in terms of capacity and delay. The values of parameters used are defined as follows: $\varepsilon_l = \varepsilon_h = 10$, $N_l = 40$, $N_h = 5$, $N_s = 15$, $\nu_l = \nu_h = 0.1$, $T_{HA} = T_{IDE} = 20$ min, $P_{BU} = 0.008$ msec, $P_{BR} = 0.001$ msec, $P_{PU} = 0.002$ msec, and $P_{PR} = 0.005$ msec.

The average subnet residence time is the expected duration that an MN stay in a subnet. Hence, as the average residence time increases, the MN performs less movement; then, the average handoff latency cost decreases for all schemes as shown in Fig. 5. However, our proposal outperforms MIPv6 and HMIPv6. The handoff latency cost gain of IISA over MIPv6 and HMIPv6 are 40% and

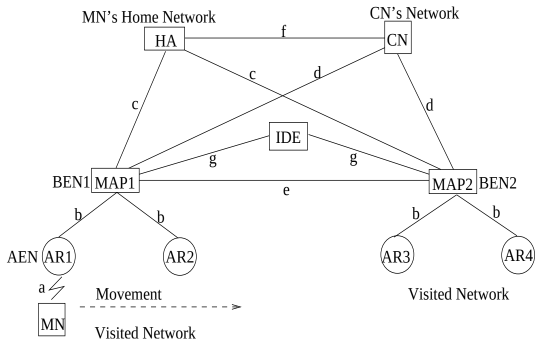

Figure 4. Network topology used for analysis.

27%, respectively. Since packet loss is proportional to the handoff latency, a similar behavior will be observed when comparing packet loss rate with residence time. Fig. 6 shows the relation between handoff latency cost

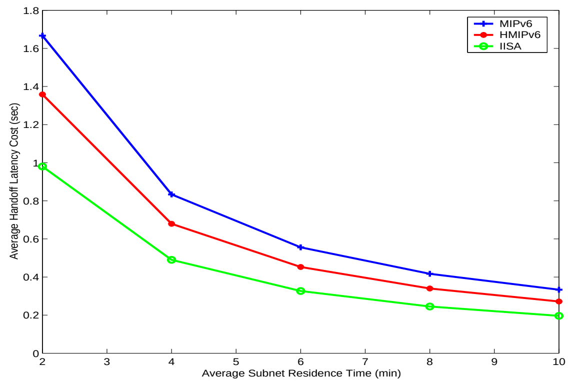

Figure 5. Handoff latency cost vs. average residence time.

and the domain crossing probability (p). We can see that the handoff latency cost of IISA is always smaller than that of MIPv6 and HMIPv6. The lower handoff latency cost of IISA results in less network signaling traffic in comparison to MIPv6 and HMIPv6. In MIPv6, the handoff latency cost remains constant since it does not differentiate intra-domain and inter-domain movement.

Fig. 7 shows the impact of the number of low-tier networks on the processing load for different values of the MN's average speed. Results show that the IDE processing load is lower than at the HA required for MIPv6. Thus, the IDE load due to intersystem and/or inter-domain handoffs is limited. On the other hand, one HA is usually used to handle MIPv6 handoff in service coverage area (e.g., one city) by network operators. We can thus conclude that a single IDE will be sufficient

TABLE I.
PERFORMANCE ANALYSIS PARAMETERS.

Parameters	Symbols	Values
Control packet size	s_c	96 bytes
Data packet size	s_d	200 bytes
Packet arrival rate	λ_p	10 packets/s
MN average speed	v_l, v_h	5.6 Km/h
Low-tier subnet radius	R_l	50 m
High-tier subnet radius	R_s	1000 m
User density in high/lower-tier networks	ρ_h, ρ_l	0.002 m^{-2}

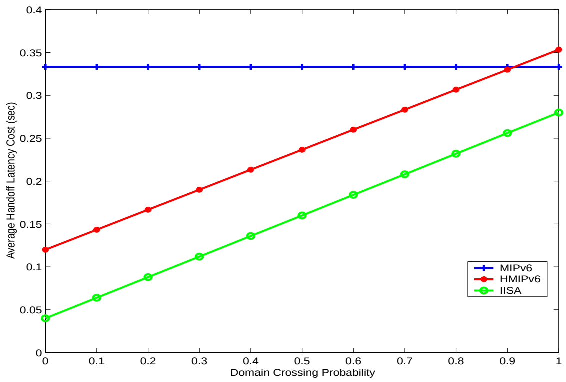

Figure 6. Handoff latency cost vs. boundary crossing probability.

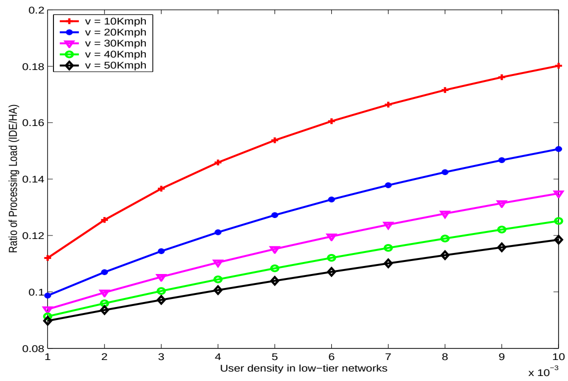

Figure 8. Ratio of processing load vs. user density.

to handle intersystem and/or inter-domain handoffs for a coverage area of one city. Fig. 8 illustrates that, as

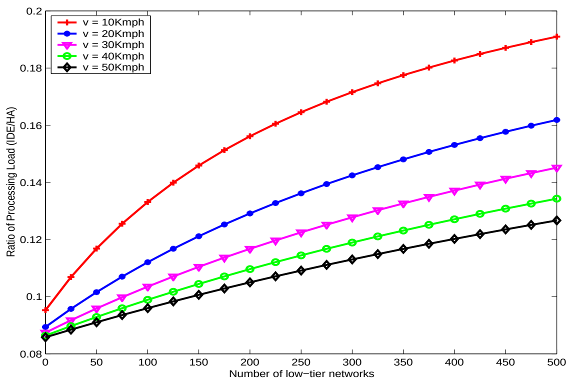

Figure 7. Processing load ratio vs. number of low-tier networks.

the user density increases, the processing load for inter-system and/or inter-domain handoffs at the IDE remains insignificant compared to the processing load at the HA for MIPv6. Fig. 9 shows that the IDE processing load increases as the number of cities increases. This means that the IDE load increases proportionally to the size of the service coverage area. Therefore, an MN with a higher average velocity is associated with a greater domain crossing rate, which results into a higher number of handoff requests. Such results encourage the deployment of the IDE through hierarchical architecture to allow the integration and the interworking of various networks.

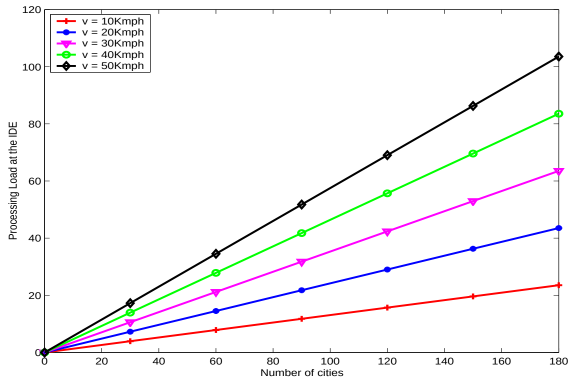

Figure 9. Processing load at the IDE vs. number of cities.

D. *Characterization of IISA*

As stated in Section II, a novel network architecture should satisfy some requirements and design goals. The IISA architecture achieves those requirements. In fact, only one new entity is added in the existing network infrastructures to allow roaming across heterogeneous networks. Other logical functionalities are implemented in existing network nodes. This allows reduction of deployment cost; hence, IISA is economic. Assume that there is O operators. The number of SLAs required to realize a roaming among all networks deployed with traditional interworking architecture is $\frac{O(O-1)}{2}$. Whereas the number of SLAs required with the IISA architecture is O. When O is very high, the IISA allows a significant reduction of the number of SLAs.

With the IDE, the IISA is an open framework, so it can support any number of networks and service providers. The IP technology is used as a common interconnection layer for next-generation networks to allow transparency of heterogeneous access technologies. With extension performed in different network nodes, roaming may be management by traditional IPv6-based mobility schemes or by our proposed handoff management scheme. This proposed handoff scheme is able to allow seamless mobility and service continuity. An equivalent level of security as provided by existing wireless is achieved in the IISA framework. It is simple enough, thus its deployment will not require strong effort and extensive costs.

VI. CONCLUSION

The integration and interworking of the existing wireless networks is one of the paths toward NGWN/4G design. Thus, several integrated and interworking architectures have been proposed in the literature. However, they cannot fulfill all requirements for real-time applications. This paper proposes a novel interworking architecture, called *Integrated InterSystem Architecture* (IISA), to enable seamless roaming across heterogeneous IPv6-based wireless environments. The IISA guarantees a seamless service continuity and alleviates service disruption during handoff as required in NGWN/4G for real-time applications. Moreover, the IISA has several advantages such as scalability, security, easy deployment and is economic.

Performance evaluation demonstrates significant gains for quality of service (QoS). Hence, we can argue that the major benefits of our proposal are minimization of handoff latency consequently packet loss, handoff failure and limited network signaling traffic overhead. In other words, by combining the IISA with the proposed handoff management scheme, it is possible to guarantee seamless roaming and services continuity across various heterogeneous IP-based wireless access networks for mobile users.

REFERENCES

[1] S. Y. Hui and K. H. Yeung, "Challenges in the Migration to 4G Mobile Systems," *IEEE Commun. Mag.*, vol. 41, no. 12, Dec. 2003, pp. 54-59.

[2] I. F. Akyildiz, S. Mohanty and J. Xie, "A Ubiquitous Mobile Communication Architecture for Next-Generation Heterogeneous Wireless Systems," *IEEE Commun. Mag.*, vol. 43, no. 6, June 2005, pp. S29-S36.

[3] 3GPP TS, "3GPP System to WLAN Interworking; System Description (Release 6)," 3GPP TS 23.234 v6.3.0, Mar. 2004.

[4] 3GPP2 TS, "cdma2000-WLAN Interworking; Stage 1 Requirements," 3GPP2 S.R0087-A v1.0, Feb. 2006.

[5] J. Y. Song, S. W. Lee, D. H. Cho, "Hybrid Coupling Scheme for UMTS and Wireless LAN interworkings," *Proc. of IEEE Vehicular Technology Conf.* (VTC'03), vol. 4, pp. 2247-2251, Oct. 2003.

[6] M. Jaseemuddin, "An Architecture for Integrating UMTS and 802.11 WLAN Networks," *Proc. of IEEE Symp. on Comput. and Commun.* (ISCC'03), Antalya, Turkey, pp. 716-723, 2003.

[7] I. F. Akyildiz, J. Xie and S. Mohanty, "A Survey of Mobility Management in Next-Generation All-IP-Based Wireless Systems," *IEEE Wireless Commun.*, vol. 11, no. 4, Aug. 2004, pp. 16-28.

[8] A. D. Assouma, R. Beaubrun and S. Pierre, "Mobility Management in Heterogeneous Wireless Networks," *IEEE Jour. of Select. Areas in Commun.*, vol. 24, no. 3, March 2006, pp. 638-648.

[9] M. Buddhikot, G. Chandranmenon, S. Han, Y. W. Lee, S. Miller and L. Salgarelli, "Integration of 802.11 and Third-Generation Wireless Data Networks", *Proc. of IEEE INFOCOM*, vol. 1, April 2003, pp. 503-512.

[10] E. Gustafsson and A. Jonsson, "Always Best Connected," *IEEE Wireless Commun.*, vol. 10, no. 1, Feb. 2003, pp. 49-55.

[11] D. B. Johnson, C. E. Perkins and J. Arkko, "Mobility Support in IPv6," IEFT RFC 3775, June 2004.

[12] H. Soliman, C. Castelluccia, K. El-Malki and L. Bellier, "Hierarchical Mobile IPv6 Mobility Management (HMIPv6)," IETF RFC 4140, Aug. 2005.

[13] GSM Association, "Inter-PLMN Backbone Guidelines," IR 34, v.3.5.2, Aug. 2004.

[14] V. G. Gupta and D. Johnston, "A Generalized Model for Link Layer Triggers," *IEEE 802.21 Media Independent Handoff Working Group*, March 2004.

[15] C. Makaya and S. Pierre, "Handoff Protocol for Heterogeneous All-IP-based Wireless Networks," in *Proc. of IEEE Canadian Conf. on Electrical and Computer Engineering* (CCECE'06), Ottawa, Canada, May 2006.

[16] M. Leibsch, A. Singh, H. Chaskar, D. Funato and E. Shim, "Candidate Access Router Discovery (CARD)," IETF RFC 4066, July 2005.

[17] J. Loughney, M. Nakhjiri, C. Perkins and R. Koodli, "Context Transfer Protocol (CXTP)," IETF RFC 4067, July 2005.

[18] M. Stemm and R. H. Katz, "Vertical Handoffs in Wireless Overlay Networks," *ACM Mobile Net. and Apps.* (MONET), vol. 3, no. 4, Dec. 1998, pp. 335-350.

Christian Makaya received the Ph.D. degree in Computer Engineering (2007) from the Ecole Polytechnique de Montreal, Canada and the M.Sc. degree in Telecommunications (2004) from the INRS-EMT, University of Quebec, Montreal and the M.Sc. degree in Computer Science (2003) from the University of Montreal, Canada. He was a Graduate Research Assistant at the Mobile Computing and Networking Research Laboratory (LARIM) and a Lecturer at the Department of Computer Engineering of the Ecole Polytechnique de Montreal.

During his doctoral studies, he was an Intern Research Engineer at Ericsson Research Canada. The current focus of his research interests is on radio resource management, interworking architectures design, mobility management, IP mobility, performance evaluation, mobile computing, service engineering and quality of service (QoS) provisioning for next-generation wireless networks.

Samuel Pierre received the B.Eng. degree in Civil Engineering in 1981 from the Ecole Polytechnique de Montreal, Québec, Canada, the B.Sc. and M.Sc. degrees in Mathematics and Computer Science in 1984 and 1985, respectively, from the UQAM, Montreal, the M.Sc. degree in Economics in 1987 from the University of Montreal, and the Ph.D. degree in Electrical Engineering in 1991 from the Ecole Polytechnique de Montreal.

He is currently a Professor of Computer Engineering at the Ecole Polytechnique de Montreal, where he is the Director of the Mobile Computing and Networking Research Laboratory (LARIM) and the Chairholder of the NSERC/Ericsson Chair in Next-Generation and Mobile Networking Systems. His research interests include wireline and wireless networks, mobile computing, artificial intelligence, and telelearning.