Context transfer for seamless micro-mobility

Juan M. Oyoqui and J. Antonio Garcia-Macias
{joyoqui,jagm}@cicese.mx
Computer Science Dept., CICESE Research Center
Ensenada, Baja California, Mexico

Abstract

Wireless networks and mobile computing technologies are having a profound impact on IP-based networks, as their basic protocols were designed without assuming mobility of the network nodes. More recently, protocols like Mobile IP have been designed to handle global mobility, and micro-mobility protocols have been proposed for the management of intra-domain mobility. However, these protocols do not address important issues required for seamless mobility, such as candidate access router discovery, target access router determination, and context transfer. In this paper, we identify the problems related to seamless mobility, underline design considerations to be observed when designing seamless mobility solutions, and propose an architecture that can be used as a framework for the implementation of these solutions.

1. Introduction

Currently, computer networks face important challenges due to the fact that in the early stages of networking all devices in the network were considered to be fixed and communication protocols were designed accordingly; nowadays mobile devices are increasingly common and design considerations have to be re-evaluated. The suite of internet protocols (IP) is not exempt to this re-evaluation and several problems need to be resolved. One of such problems is related to the fact that the IP address of a node is used to determine its location, so if the node moves within the network it must then change its IP address to be properly located. Mobile IP [13] has dealt with this problem by allowing nodes to change their IP point of attachment (IP-POA) and their IP address within the network. However, Mobile IP has several drawbacks when several handoffs1 take place within a domain; this has led to the study of micro-mobility. Thus, usually two types of protocols are used to handle mobility: Mobile IP is used to handle mobility between domains, and a micro-mobility protocol takes care of intra-domain mobility. Among the most important micro-mobility protocols we can cite Hierarchical Mobile IP, Cellular IP, and HAWAII, which will be briefly described later. An interesting problem posed by mobility is context transfer, which arises with the need to minimize the time needed to associate a mobile node (MN) with its new access router. This process is achieved by transferring information about the current state of a MN, known as context, from its previous to its new access router. Thus, an important goal in context transfer is to quickly re-establish in the new access router the context it had in its previous access router; of course, this should be faster, or at least equal, to the case of establishing a new context from the beginning. Context transfer involves deciding which of the potential access routers should be chosen to service the MN; such decision can be based not only on received signal strength but also on the capabilities of the routers and how appropriately they can satisfy needs of the MN.

This paper is organized as follows: section 2 introduces the problem of mobility in IP-based networks; emphasis is made on the differences between global mobility, as handled by Mobile IP, an intra-domain mobility, which requires specialized micro-mobility protocols (some of which are presented). Section 3 discusses why seamless mobility requires more than just a mobility protocol, as it involves discovering all access routers that can be candidates for becoming the target access router for completing a handoff, some criteria and evaluation methods for selecting the appropriate target access router, and mechanisms for performing (or attempting to perform) context transfer in order to minimize the impact of handoffs while moving. In section 4 we identify the problems related to seamless mobility, highlighting design considerations that are crucial for the conception of seamless mobility solutions including protocols for candidate access router discovery and context transfer; observing these considerations, we illustrate how our proposed architecture can be used as a framework for the implementation of these solutions. We conclude in section 5 by summarizing our presentation and then indicate open issues, the scope of our work, and trace the direction for future developments.

1 The terms handoff and handover are used interchangeably in the literature.
2. Mobility in IP networks

As Reinbold notes [16], there are several important issues for managing mobility in IP networks:

- **Handover management**: It concerns the management of the changes of IPPOA of the MNs as they move across the wireless network.

- **Passive connectivity and paging**: Mobile devices typically have limited power capacity and transmissions should be kept to a minimum to avoid battery exhaustion. However, if a MN does not emit any signals when it changes its IPPOA, it will be impossible to forward incoming packets as its location will be unknown. A technique known as IP paging has been under development recently to deal with these issues [15, 2].

- **Quality of service (QoS)**: As application demands grow due to ever-increasing multimedia traffic, networks should be able to provide guarantees in terms of assured bandwidth, maximum delays and jitter, etc. The same quality of service should be kept as MNs roam from cell to cell performing handoffs.

- **Intra-domain traffic**: An important part of the traffic in wireless networks takes place between MNs in the same administrative domain. This intra-domain traffic is more dynamic than inter-domain traffic and should be efficiently supported.

Currently, the most popular solution for mobility management in IP networks is Mobile IP [13]. In principle, Mobile IP can handle both global and local mobility. However, it requires that the mobile’s home network be notified of every change in location. Moreover, route optimization enhancements further require that every new location be registered with hosts that are actively communicating with the mobile node. All these location updates incur communications latency and also add traffic to the wide-area portion of the internetwork. Therefore, Mobile IP is not efficient when there are large numbers of portable devices moving frequently between small cells. It has also been demonstrated that, when used for micro-mobility support, Mobile IP incurs disruption to user traffic during handoff, and high overhead due to frequent notifications to the home agent [12]. Another type of protocol, a micro-mobility protocol [16, 1], is then needed for local environments where mobile hosts change their point of attachment to the network so frequently that the basic Mobile IP tunneling mechanism introduces network overhead in terms of increased delay, packet loss and signaling.

Some of the most prominent micro-mobility protocols are:

**Hierarchical Mobile IP**: Hierarchical Mobile IP [3] was designed as a natural extension to Mobile IP to provide efficient support for micro-mobility. The basic idea is that while the MN is moving within a domain, the mobility agent at its home network does not need to be aware of that motion. Once in a foreign network the MN performs a regional registration with a Gateway Foreign Agent (GFA) that will be aware of the local mobility but will not inform it to agents outside the domain.

**Cellular IP**: A goal of Cellular IP [19] is to integrate principles of cellular telephony in IP networks. A Cellular IP network is connected to the Internet via a gateway router. Mobility between gateways (i.e., Cellular IP access networks) is managed by Mobile IP, while mobility within access networks is handled by Cellular IP. A Cellular IP node constitutes the universal component of a Cellular IP network, since it serves as a wireless access point but at the same time routes IP packets and integrates cellular control functionality traditionally found in Mobile Switching Centers (MSC) and Base Station Controllers (BSC). Cellular IP nodes are modified IP nodes where standard routing is replaced by Cellular IP’s own routing and location management functions.

**HAWAII**: Under the approach of HAWAII [14], MNs retain their network address while moving within a domain; this way, any mobility agent in the home network —if using Mobile IP for global mobility— and any corresponding hosts, are not aware that the host has performed intra-domain mobility. When a mobile host moves within a domain (performing a handoff), all involved forwarding tables are modified to redirect packets to the mobile’s new location. These changes are made under one of four possible path setup schemes that determine when, how, and which routers are updated, namely: MSF, SSF, UNF, and MNF.

3. Seamless mobility in local networks

As seen before, mobility in geographically restricted networks (e.g., intra-domain mobility) is different to global mobility as routing paths between a MN and the network may change frequently and rapidly. Sometimes, the MN establishes certain services in its current location on the network, and as it moves performing handoffs, it wishes to benefit from the same services at its new location. Establishing these services initially may require a considerable amount of time for setting up a certain state within the network and to carry out certain protocol exchanges. When it moves, if the MN is required to re-establish those services by the same process as it uses to initially establish them, delay-sensitive real-time traffic may be seriously impacted. An alternative is to transfer enough information on the context transfer-candidate service state, or context, to the new
location so that the services can be re-established quickly, rather than require the mobile host to establish them from scratch.

3.1 Issues in context transfer

The Internet RFC 3374 document [8] published by the SeaMoby working group [6] provides the following terminology that helps to better understand concepts related to context transfer:

**Context.**- The information on the current state of a service required to re-establish the service on a new subnet without having to perform the entire protocol exchange with the mobile host from scratch.

**Context Transfer.**- The movement of context from one router or other network entity to another as a means of re-establishing specific services on a new subnet or collection of subnets.

**Context Transfer Candidate Service.**- A service that is a candidate for context transfer. Services that are concerned with the forwarding treatment of packets, such as QoS and security, or involve granting or denying the mobile host access to the network, such as Authentication, Authorization, and Accounting (AAA), are considered to be context transfer-candidate services.

There are two basic motivations for context transfer: first, the need to quickly re-establish context transfer-candidate services without requiring the MN to explicitly perform all protocol flows for those services from scratch. An additional motivation is to provide an interoperable solution that works for any Layer 2 radio access technology.

There are certain limitations that can make context transfer a sub-optimal solution for re-establishing context transfer-candidate services. Such limitations are [8]:

- **Router compatibility.**- Context transfer between two routers is possible only if the receiving router supports the same context transfer-candidate services as the sending router. This does not mean that the two nodes are identical in their implementation, nor does it even imply that they must have identical capabilities. A router that cannot make use of received context should refuse the transfer. This results in a situation no different than a mobile host handover without context transfer, and should not be considered an error or failure situation.

- **Requirement to re-initialize a service from scratch.**- The primary motivation for context transfer assumes that quickly re-establishing the same level of context transfer-candidate service on the new subnet is desirable. And yet, there may be situations where either the device or the access network would prefer to re-establish or re-negotiate the level of service. For example, if the mobile host crosses administrative domains where the operational policies change, negotiation of a different level of service may be required.

- **Suitability for the particular service.**- Context transfer assumes that it is faster to establish the service by context transfer rather than from scratch. This may not be true for certain types of service, for example, multicast, “push” information services.

These limitations should be taken into account in the design considerations of seamless mobility solutions. It should also be remembered that context transfer aims for an improvement in handover performance by sustaining the transfer-candidate services provided to a MN as it moves. Thus, any possible solution must provide performance that is equal to or better than re-initializing the context transfer-candidate service between the mobile host and the network from scratch. Security considerations should also be observed, including information privacy, transfer legitimacy, and security preservation.

3.2 Candidate access router discovery

Before performing a context transfer and a handover takes place, the MN must first identify which potential new access routers there are, and which one of them is the most appropriate to be associated with after the handover. Discovering candidate access routers (CARs) involves identifying the IP addresses of the CARs and finding the capabilities of those CARs. Once all possible CARs are discovered, the MN will then choose one of them as the target access router (TAR) for completing the handover. Such decision involves obtaining information about the CARs so the best decision about the TAR can be made. There are some definitions
that should be presented in order to better understand the problem of CAR discovery and selecting a TAR (refer to Figure 1):

**Access Point (AP).** A radio transceiver by which a MN obtains Layer 2 connectivity with the wired network.

**Access Router (AR).** An IP router residing in an access network and connected to one or more APs. An AR offers IP connectivity to MNs.

**Candidate AR (CAR).** An AR to which a MN has a choice of performing IP-level handover. This implies that the MN has the right radio interface to connect to an AP that is served by this AR, as well as the coverage of this AR overlaps with that of the AR to which the MN is currently attached to.

**Target AR (TAR).** An AR with which the procedures for the MN’s IP-level handover are initiated. The TAR is selected after running a TAR Selection algorithm that takes into account the capabilities of CARs, preferences of the MN and any other local policies. It is then clear that there are two basic problems related to CAR discovery: first, mapping from a Layer 2 identifier (e.g., a MAC address) for an AP to the IP address of the CAR; and second, identifying the capabilities of the CAR.

### 4. An architecture for seamless micro-mobility

As we have seen, intra-domain mobility is handled by a micro-mobility protocol that allows MNs to handover from cell to cell in a wireless network. Although several micro-mobility protocols have been proposed and tested (some of which were discussed earlier), they all assume that a TAR has already been selected from a number of CARs. Also, when performing a handoff, it would be beneficial (if not mandatory) for the MN to find at the TAR the same context it had at its previous AR.

Discovering the possible CARs, selecting a TAR from them, and performing context transfer are open issues that we address in this paper. It is worth noting that these issues are being discussed in the IETF Seamoby Working Group [6] and some problem specifications [8, 9, 17, 18], and early proposals [10, 11], still considered work in progress, have arisen.

We identify the following actions that take place when a MN performs a handoff involving CAR discovery and context transfer (see Figure 2):

1. A layer 2 event takes place, which triggers the need for a handoff.
2. The CAR discovery process is started and a TAR is selected.
3. Handoff to the TAR starts, involving layers 2 and 3 actions.

As indicated, a layer 2 event can be interpreted as a warning on the convenience (or imminency) of starting a handoff. We emphasize that this is a layer 2 event, so the MN does not have direct layer 3 communication with any of the CARs, as it does not know their IP address. This layer 2 event can be, for instance: establishing a layer 2 (wireless) link with a new AR, the MN receives a beacon from a CAR, or degradation of the quality of the signal that the MN receives from its current AR.

Once the layer 2 trigger appears, the possible CARs should be discovered in order to determine a target for the handoff. If the MN receives (layer 2) signals from several ARs, the problem now lies in: determining their IP addresses and then discovering their capabilities. This process is called candidate access router discovery (CARD). After that, a TAR should be selected based on which CAR best fits the needs of the MN. The protocol for CAR discovery and TAR selection should contemplate the following functionalities:

- **Reverse address translation.** Involves mapping the layer 2 identifier of a new AP, connected to the CAR, to the IP address of the CAR to allow communication with the MN. In cases when the MN can have IP connectivity with the CAR before making a handoff decision, this functionality is carried out trivially. However, if the MN can only listen to the layer 2 identifiers of the new APs before making a handoff decision, it is necessary to have a mechanism to perform the reverse address translation. Some possible solutions would be to use information stored in the cache of the MN, or...
alternatively to ask the current AR to try to map the layer 2 identifier of the CAR to its IP address.

- **Discovery of capabilities.** Information about the capabilities of the CARs can help the MN to make appropriate handoff decisions; it can be particularly useful as input for a procedure to select a TAR. Some of the capabilities can be static, others could change at large temporal intervals, while some others could change very dynamically. A "time to live" or expiration time can be assigned to the capabilities and the AR can assist the MN to obtain the capabilities of the CARs.

The discovered capabilities will help in making handoff decisions, so we propose to use concepts related to Policy Based Network Management (PBNM). Furthermore, we propose to leverage the work of the IETF Policy Framework Working Group [7], which defines an object-oriented model [4] to represent information about policies. This model defines a two-level hierarchy for object classes: structural classes representing policy information and control, and association classes indicating how many instances of the structural classes are inter-related. This working group suggests the use of LDAP [5] as the protocol for accessing policy databases.

The following scenario illustrates the different entities involved in a handoff with context transfer and the interactions between them, as specified by our seamless mobility architecture (the related sequence diagram is shown in Figure 3). When a MN is moving outward the periphery of its current wireless cell, an event —for instance, degradation of the quality of the signal received from its current AP— makes the MN listen to layer 2 advertisements that other APs within reach are sending (e.g., beacons). These advertisements contain the layer 2 (MAC) address of the candidate APs, and the MN collects a list of these addresses; the list is then sent to its Old AR soliciting that it finds, for each AP, the IP address of the AR associated with it, as well as discovering the capabilities of the AR. In order to do this, we propose the use of a management repository —possibly located at the edge or gateway access router (GAR), at the top of the hierarchy of the administrative domain. When a device joins the network, its associated AR will inform the repository about the device’s capabilities; furthermore, each AR will be charged of updating the repository with the general status of the wireless cell it manages, whenever a change in the cell takes place and modifies the status of the current load, available bandwidth, etc. This way, the repository will always have an updated “snapshot” of the status of the whole domain, and will also have the updated management policies for the domain. This information will be periodically advertised down the hierarchy to the ARs. Thus, when the Old AR wishes to know the capabilities of the CARs in behalf of a MN, it will just check the information it has about those ARs, or could just simply consult the repository. The information gathered from the repository will help to determine which of the ARs is best suited to become the TAR. Then, the handoff is performed and communication at the IP level is established between the MN and the TAR. There are certainly security considerations to be observed when designing the CARD protocol. However, security issues are outside the scope of this paper and are somehow orthogonal to the core problems we focus on.

### 4.1 Protocol design considerations

In order to be efficient, the CARD protocol should be able to send messages along the IP signaling of the handover; we are exploring the use of in-band signaling\(^5\) for this purpose. Also, to optimize execution time, it should be possible to send signaling messages independently of the IP flow, allowing for the use of a standalone CARD protocol; one way to achieve this is through the use of ICMP messages.

Regarding the context transfer protocol, it should be recalled that the main objectives are: to reduce latency, packet loss, and avoid re-initialization of signaling between the MN and the TAR. The proposed protocol should include the following:

- A format for representing features related to contexts.
- Auxiliary messages for context transfer, including initiation, authorization, and informing the MN of the status of the transfer.

\(^{5}\)With this approach, signaling information can be coded inside unused IP header fields, for example: IP options
• Messages to carry out the actual context transfer, normally before the handover, but also allowing the transfer after or during the handover.

Network considerations, such as congestion control, should be observed. ICMP and UDP do not provide congestion control, while TCP/SCTP do provide it. Congestion control can slow down context transfer. A related issue is error correction, provided by TCP and SCTP via mandatory checksums, while UDP provides only and optional checksum. It is worth noting that recent work in progress within the IETF has resulted in early proposals for a context transfer protocol, although no decision has been made concerning the transport method for the protocol [11]. We are concurrently evaluating the advantages and disadvantages of the different options. Independently of the transport method, a protocol for context transfer should be able to report when context transfer can not be achieved, and this impossibility should not negatively affect the network or ongoing user sessions.

5. Summary and Conclusions

In this paper we have discussed the problems that arise with IP-based mobility and have presented the approaches and protocols proposed for managing global and micro-mobility. We have evidenced how these proposals do not address important issues required for seamless mobility, such as candidate access router discovery, target access router determination, and context transfer. We have identified the problems related to seamless mobility, underlined design considerations to be observed when designing seamless mobility solutions, and have proposed an architecture that can be used as a framework for the implementation of these solutions.

The presented work will serve as a basis for future developments, and for continuing work in still open issues, as we discuss next. The design considerations we have proposed will be crucial for the implementation of prototypes for our proposed CARD and context transfer protocols, which constitute (near-)future work. However, we should note that the scope of our work does not reaches into the security domain, as we focus in functional considerations. Some observations can be made regarding related security issues: before performing context transfer, the MN should be properly authenticated and authorized; communication between the ARs should also be made through secure channels, but these should be established before performing context transfer to avoid incurring additional latency to the process. These considerations also apply to the CARD protocol.

We have proposed to use a Policy Based Network Management approach for the problem of TAR selection. We are currently working on the definition of policies that can be used as inputs for the algorithm that decides which one of the CARs is best suited to become the TAR; this involves storing the policies, propagating them through the domain, informing when changes to the policy base are made. Error and exception handling is a critical issue, for example, what happens if context transfer is requested but can not be carried out? Finally, we should underline that we focus on intra-domain seamless mobility, but our proposals for CARD and context transfer could be generalized to the case of inter-domain mobility, although that case lies outside the scope of our current research efforts.

Acknowledgments

The work of Juan M. Oyoqui has been supported by a scholarship granted by the Mexican Science and Technology Council (CONACyT).

References


