Seamless Support of Multimedia Distributed Applications Through a Cloud

Stefano Ferretti, Vittorio Ghini, Fabio Panzieri, Elisa Turrini
Department of Computer Science, University of Bologna
{sferrett, ghini, panzieri, turrini}@cs.unibo.it

Abstract

We describe a cross-layer architecture we are developing in order to offer mobility support to wireless devices executing multimedia applications which require seamless communications. This architecture is based on the use of pairs of proxies, which enable the adaptive and concurrent use of different network interfaces during the communications. A cloud computing environment is used as the infrastructure to set up (and release) dynamically the proxies on the server-side, in accordance with the pay-as-you-go principle of cloud based services.

1. Introduction

In this document we describe an on-going project aimed at the provision of support to mobile devices, equipped with multiple wireless network interfaces (NICs), running multimedia applications. Specifically, we have developed a cross-layer architecture which offers *always-connected* services by exploiting all the networks available to the user, and by dynamically adapting their use on the basis of their performance and costs. Thus, our architecture can provide its users with reliable communications even in presence of vertical and horizontal handoffs, and can optimize those communications, augmenting the probability of meeting application QoS requirements such as responsiveness, availability, continuity of the communication service.

Our cross-layer architecture is able to manage different types of communication flows, based on either the UDP or the TCP transport protocols. Basically, it offers an adaptive use of different network technologies, transparently to the applications, by exploiting a pair of proxies for each mobile device. A client proxy needs to be installed (or simply loaded in memory) on the mobile device, while a server proxy executes over the Internet on a fixed host with a public IP address(see Figure 1).

The cross-layer part of the proposed architecture is confined into the client proxy which is the only component that depends on the mobile device operating system. In essence, the client proxy is in charge of managing the wireless networks and dynamically deciding which network to use on a packet basis. The server proxy is aware that each client can use multiple nets, and that messages from the same client (i.e., belonging to the same communication flow) may be received coming from different IP addresses. Hence, once identified all the messages that belong to the same client, it rebuilds the communication flow from that client and forwards that client's messages to their destination. Thus, the server proxy appears as the client to the destination host. Moreover, the pair of proxies implements packet load balancing and retransmission on the wireless channels to meet application QoS requirements.

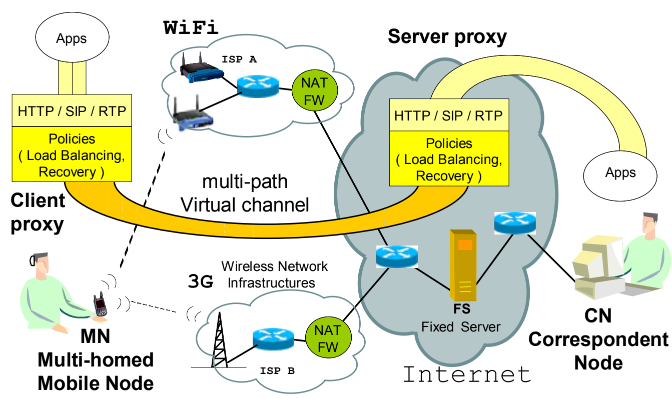

Figure 1. Cross-layer mobility management architecture

Our mobility management mechanism overcomes a number of limitations of other existing approaches. In particular, the Mobile IPv6-based architectures, such as Fast Handover MIPv6 [1] and Proxy MIPv6 [2], need that network infrastructures be modified to add IPv6 capabilities and, moreover, do not allow the simultaneous use of the multiple network interfaces of the mobile node. Others solutions at the ISO/OSI transport layers, such as Datagram Congestion Control protocol [3] and Mobile Stream Control Transmission protocol [4] require the end user applications be modified. In contrast, our solution does not require modifications of the access networks or the applications that may use "proxies", such as Web-based and SIP/RTP-based multimedia applications.

2. The cloud-based architecture

The basic architecture composed of the two proxies, in Figure 1, has been fully implemented and has already been exploited to support, in an urban context, different types of applications, such as Web 2.0 applications [5], online games, Voice over IP applications [6], mobile video surveillance systems [7].

However, a practical limitation of our approach is that a server-proxy is needed for each corresponding client, even if a given server-proxy may serve more than one client simultaneously. In addition, the server-proxy should be located close to the client so as to reduce the network latency and optimize the server's responsiveness. Hence, in order to deploy effectively our mobility management mechanism over the Internet and make it work on a large-scale, some scalable solution is needed. This is where the cloud computing enters into the picture. Cloud architectures support dynamic resource provision and allocation so as to allow applications to scale on demand, and minimize the waste associated with underutilized computing resources. The idea is that our server proxies might be dynamically allocated (released) once a client joins (leaves) the network. The client might contact a gateway that allocates the resources and provides details on the allocated server proxy. This would augment the scalability of the system and optimize the resource usage. It is worth pointing out that also the gateways run on top of the cloud subsystem.

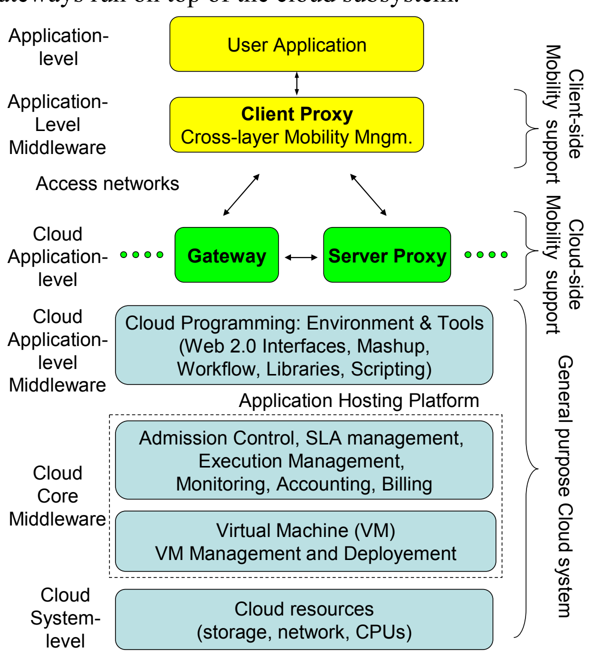

Figure 2. System architecture

In a Infrastructure as a Service (IaaS) cloud, i.e. a system that provides low-level services such as virtual machines which can be booted with a previously defined hard-disk (and software) image [8], a virtual machine executing our server proxy can be set up each time a (new) mobile client makes a request. It is worth observing that our proposal differs from typical Web-based cloud applications as it requires that a client proxy be installed and executed on the client side.

The overall system architecture we propose is composed of the following three principal subsystems that are depicted in Figure 2:

i) the general-purpose cloud subsystem that virtualizes the resources and provides the typical cloud services necessary to implement the server side of the distributed terminal mobility support subsystem. In particular this subsystem is composed of four layers: the lower three layers are in charge of, respectively, managing physical resources of the hosts, running the virtual machines on the selected host of the cloud, and providing admission control, accounting and billing. The higher layer provides libraries, tools and environment for the applications.

ii) the cloud-side mobility support subsystem, deployed on the cloud, provides the users with the application specific always-connected service, and is composed of the previously cited gateways and server proxies. It is responsibility of the gateway the selection, for each given mobile user, of the most suitable location for the server proxy, so as to reduce the network latency between the pair of proxies.

iii) the mobile user subsystem, located at the mobile device, that runs both the user application and the client proxy that hides the effects of the user movements.

3. Concluding remarks

We are planning to carry out a validation and a thorough experimental assessment of the performance of our cross-layer architecture as soon as its development will be completed. In addition, we would like to extend our study on this class of architectures to investigate the impact of dependability issues, such as fault tolerance and security, on their design.

References

[1] R. Koodli, "Fast Handover for Mobile IPv6", IETF RFC 4068, July 2005.
[2] S. Gundavalli et al., "Proxy Mobile IPv6", IETF Internet Draft, draft-ietf-netlmm-proxymip6-01.txt, June 2007.
[3] E. Kooler et al., "Datagram Congestion Control Protocol (DCCP)", IETF RFC 4340, March 2006.
[4] M. Riegel, M. Tuexen, "Mobile SCTP," IETF Internet draft, draft-riegel-tuexen-mobile-sctp-07.txt, Oct. 2006.
[5] S. Ferretti et al., "A web 2.0, location-based architecture for a seamless discovery of points of interests," proc. of AICT 2009, May 2009, Venice, Italy. IEEE Computer Society.
[6] V. Ghini et al., "Always Best Packet Switching: the Mobile VoIP Case Study", Journal of Communications (Academy Publisher), Oct. 2009.
[7] V. Ghini, G. Lodi, and F. Panzieri, "Mobile e-witness," Multimedia Tools Appl., vol. 37, no. 3, pp. 293–318, 2008.
[8] M. Fouquet et al., "Cloud computing for the masses," U-NET '09: proc. of ACM workshop on User-provided networking: challenges and opportunities. New York, NY, USA: ACM, 2009.