Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Computer Communications 29 (2006) 730–740
computer communications
www.elsevier.com/locate/comcom

Dynamic resource management for QoS provisioning over next-generation IP-based wireless networks

Sotiris Maniatis*, Eugenia Nikolouzou, Iakovos S. Venieris

School of Electrical and Computer Engineering, National Technical University of Athens, 9 Iroon Polytechniou Street, 157 73 Athens, Greece

Abstract

A significant issue in current research pursuits is the formulation of the requirements and basic design options for the next-generation wireless network architecture. The next-generation of wireless systems will support a diverse set of access technologies and mobile devices, formulating a broad heterogeneous environment with increased requirements on network support operations. It is expected that the demanding breed of multimedia applications will even more considerably require Quality of Service support throughout the end-to-end path. This paper first provides a tutorial approach on next-generation wireless network architectures and more specifically on end-to-end QoS provision. We claim that dynamic resource management in the Core Network is a necessity due to the increased heterogeneity of the new environment. We subsequently present our proposal regarding a dynamic resource management scheme that is based on the concept of the Resource Pools. The Resource Pool concept is deeply analysed within the paper and simulation results prove its correctness and appropriateness.

Keywords: Next-generation wireless networks; Quality of service; Dynamic resource management; Network services

1. Introduction

The next-generation wireless networks will be principally formed through the evolution and convergence of current mobile communication systems and the IP technology. The foreseen architecture will retain the well-known bi-level structure, consisting of a multi-domain Core Network (CN) offering IP connectivity and services, and a set of wired and wireless Access Networks (AN) offering the last mile connectivity services to mobile users. The convergence concept has gained significant attention during the last years in the research, industrial and standardisation community, leading to a number of convergence scenarios and propositions, like in Refs. [1–3]. The fundamental characteristic of next-generation networks is heterogeneity. Heterogeneity is the result of the different operations and capabilities of the multiple converged access networks, the reconfiguration and adaptation capabilities of end-user devices and applications, as well as the interworking model resulting from the various handover scenarios and the existence of multiple providers.

The architecture of next-generation wireless networks will therefore have to deal with heterogeneity. The latter imposes strict requirements that have to be satisfied in order to provide a smooth and seamless service. These requirements affect the basic operations of the core network, including mobility management, network resource and Quality of Server (QoS) management, and overall AAA operation control, among others. This paper focuses on network resource management for QoS provisioning in a heterogeneous next-generation wireless network environment. The main motivation of our work is that the heterogeneous nature of such environment necessitates the existence of a dynamic resource management layer in the core network of the architecture, as briefly explained hereafter.

In third generation (3G) or previous systems, the CN is not considered the bottleneck of the overall network. In such networks the bottleneck is always the pure wireless part of the AN, allowing the operators to properly dimension the remainder network (the wired part of the AN as well as the CN) so that minimum congestion will occur in that. Dimensioning here does not necessarily mean

* Corresponding author. Tel.: +30 210 772 2318; fax: +30 210 772 1092.

E-mail addresses: sotos@telecom.ntua.gr (S. Maniatis), enik@telecom.ntua.gr (E. Nikolouzou), ivenieri@cs.ntua.gr (I.S. Venieris).

0140-3664/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2005.07.016

over-provisioning as one could claim, but instead a more or less static management of CN resources, according to predicted traffic patterns, and based on the standard CN's QoS capabilities. This is eventually true for any homogeneous wireless network that gives to the administrator the opportunity to study the traffic patterns, predict the traffic demands and at last appropriately dimension the network.

Dimensioning in homogeneous wireless networks, although possible, is indeed very demanding and difficult, mainly due to user mobility. Dimensioning is, however, far more difficult for heterogeneous networks that expose dissimilar traffic patterns stemming from the multiple available access networks. Moreover, user mobility does not only entail an intra-system handover, but it may involve an inter-system one, resulting in considerable adaptation of the traffic involved. Furthermore, with the advent of ad hoc wireless networks, which can be configured on demand, dimensioning and provisioning of the core network becomes even more intense.

Due to the aforementioned facts, dynamic resource management in the core of next-generation wireless networks appears to be a necessity. There is common consensus on the fact that CN will be a pure IP-based network, meaning that all CN operations will be based on mechanisms and protocols developed and used in the IP world. The chief standardisation body regarding IP is the Internet Engineering Task Force (IETF) [4]. Concentrating on QoS, which is the focal point of this paper, IETF has basically defined two frameworks: the Integrated Services (IntServ) and the Differentiated Services (DiffServ). The former offers QoS guarantees with the aid of the RSVP protocol but exposes scalability problems, while the latter provides only soft-QoS-guarantees. The introduction of the Bandwidth Broker (BB) concept by the Internet2 QBone [5] was an effort to basically cater for resource management and admission control over DiffServ networks. However, the BB architecture has not been standardized in IETF. As it will be presented in Section 2, current research trends favour the BB-enhanced DiffServ framework compared to IntServ and RSVP. Our work, presented in this paper, is also based on this framework.

The dynamic resource management presented in this paper is based on the concept of Resource Pools. Resource Pools (RPs) try to overcome the scalability problems of the centralised BB concept, by introducing a distributed and highly scalable resource management scheme, and cater for a dynamic distribution of resources to the different heterogeneous access networks connected to the CN based on real traffic demands. The fundamental idea is to organize a number of Edge Routers[1] (ERs) sharing a common bottleneck element into groups, which are called RPs. Those groups will provide a dynamic and efficient way for sharing and shifting the available resources between RPs based on real traffic demands. In the context of the RPs model, some algorithms are proposed to cater for dynamic (re)-distribution of resources among RPs, in case the real resource demand is not met by the initial provisioning scheme. Those algorithms derive from the AQUILA [6] framework, which comprises a wired network, but are further enhanced to encounter the requirements introduced by the heterogeneous wireless segments.

The rest of the paper is structured as follows. Section 2 presents a tutorial-in-nature approach for the next-generation wireless network architecture, identifying the basic network elements and operations, discussing the control layer needs for QoS provisioning and the network services provisioning issue, and lastly summarizing the state of the art and current trends in end-to-end QoS. Section 3 explains the dynamic resource management mechanism, while Section 4 presents some simulation results that prove its appropriateness. Finally, Section 5 contains the conclusions and future work.

2. Next-generation wireless network architecture

There is a common consensus about the fundamental building blocks of the next-generation wireless network architecture. A common IP-based Core Network provides the mobile devices with the basic IP connectivity and uses only native IP protocols (IETF-based) for every operation: network address assignment, network management, mobility, quality of service, and AAA, among others. The Core Network retains the functionality of the well-known 3G Core Network, but it can be considered an evolution of the latter, since the intention is to decouple the wireless access technologies from the core network serving them. The various access segments will connect to this unified core through a generic interface (similar to Iu in 3G terminology), communicating with special routers that will have the ability to perform the adaptation functions needed for the underlying access technologies. We have to stress here that the Core Network will be able to serve not only wireless access segments but also wired ones. Fig. 1 depicts the envisioned architecture of the next-generation wireless environment.

The morphology of the Core Network follows the typical IP network setup. Special Edge Routers (ERs) interconnect the various access segments to the core. Apart from performing access-specific adaptation functions, these routers operate as typical edge routers in the IP world. Note here, that instead of employing specialised ERs, an interworking unit could be provided at the edge of each access segment to perform the adaptation functions. In this case, ERs are conventional edge IP routers. The Core Network additionally consists of Core Routers (CRs) that interconnect ERs and Border Routers that provide the connectivity to external IP networks. In this sense,

[1] In the context of this paper, Edge Router is considered the network element which interconnects an access network with the CN.