Universitat Ramon Llull

TESI DOCTORAL

Títol **New Challenges in Quality of Service Control Architectures in Next Generation Networks**

Realitzada per **Àlex Vallejo Blanxart**

en el Centre **d'Enginyeria i Arquitectura La Salle**

i en el Departament **d'Informàtica**

Dirigida per **Dr. Jordi Dalmau Royo**

C. Claravall, 1-3
08022 Barcelona
Tel. 936 022 200
Fax 936 022 249
E-mail: urlsc@sec.url.es
www.url.es

Acknowledgments

It is true, a thesis is a very long and lonely journey, but not all the way. It would have been impossible to carry out this thesis without the collaboration of many people. Some of them just helped a bit at some crucial points, some have been around transversally, some we have never met in real life and two of them were even born in the course of this work.

Firstly, I would like to thank my supervisor, Dr. Jordi Dalmau for accepting me as his PhD student. His vision of the broadband networks dramatically changed my vision of the networking technologies.

The present work is the result of the collaboration with a number of researchers. I would like to thank my co-authors: J. Abella, A. Campos, X. Canaleta, D. Cutiller, J. Dalmau, C. Duz, A. Orriols, J.Ruiz, J.M. Selga, D. Vernet, G. Villegas, and specially, A. Zaballos. I would like also to thank those who made the IPv6 testbed and the QoS broker possible from the very beginning in the Maracaná Lab: A. Campos, M. Capdevila, C. Duz, J. Garcia, J. López, D. Mateo, J. Nubiola, S. Rebordosa, J. Sánchez, C. Seguí, J. Seguró and F. Valpondi. I am also very grateful to the rest of the Networking Technologies team, namely J. Dehesa, X. Cadenas and G. Corral.

I am really grateful to J. Mundet and his scientific investigation method class, from which we learned the real state of the art of the scientific community. I would like to extend this acknowledgement to M.H. Sherif for his patience, P. Dini for his kindness and A. Odedra, T. Tsou and M. Hayashi from the SG-11 at ITU-T's NGN-GSI for their invaluable collaboration.

My thanks also go to E. Villanueva for her assistance with the linguistic reviews in some parts of the thesis, and very especially, to L. Kinnear for religiously reviewing all my papers, including this thesis.

I would like to express my gratitude to Enginyeria i Arquitectura La Salle for its support. I would also like to thank the Ministerio de Educación y Ciencia, which has supported the CSI-RHET project (TEC2005-08068-C04-04/TCM), the Ministerio de Industria, Turismo y Comercio, which has supported the "Telefonia sobre IP con calidad de servicio" project (FIT-330200-2004-189) and the European Commission, which has supported the FP6-IST-OPERA project (number 026920).

Last, but not least, I would like to thank all my family and friends for their unconditional support and infinite patience. I would like to very especially thank those who suffered my busy nights and weekends.

Contents

List of figures

List of tables

1. INTRODUCTION

The efforts of data networks, and specially Next Generation Networks (NGN), have focused on providing converged global services as well as ubiquity and network transparency for users and applications. These new multimedia services along with the increasing enterprise communications passing through data networks have generated the need for mechanisms to differentiate data traffic and services, and therefore to provide Quality of Service (QoS). New architectures to meet these needs based on policy-based network management to control QoS network resources and to ensure scalability and seamless end-to-end QoS control are being developed worldwide.

This QoS control is especially critical in the Next Generation Access networks (NGAs) which are usually the traffic bottlenecks in end-to-end flows. The improving features of the Broadband Power Line networks (BPL) have led to their incorporation in NGAs. Nowadays these BPL networks provide high-speed symmetric and cost-effective alternative “last mile” broadband access. Besides, in the future, the Smart Grid is to be a definitive sustainable solution enabling distributed generation and the extension of the intelligent control over electrical power grid functions to the distribution level and beyond, and it will have to include a communications solution. These Smart Grid networks will have to manage huge quantities of real-time traffic over a heterogeneous network which may consist of wireless and BPL nodes.

Furthermore, the growing requirements of the network market for functional intelligent systems in the QoS area are currently motivating important research and development work. Intelligent systems, which can provide solutions to non-computational problems, are important in tackling practical computing problems for service providers and operators. This thesis aims to provide practical contributions to the data networks QoS field including the advantages of intelligent systems.

This chapter presents a comprehensive overview of the thesis. The framework that embraces this thesis is detailed along with the fundamentals that motivated this research and the lines of research are related to the different projects in which we have taken part. The main goals of this thesis are also described in this chapter. Finally, the organization of this thesis is set out.

1.1. Framework: The QoS needs in a changing world

Pioneers of Internet [212] could never have imagined back in the 70s they were about to change the history of humanity in a definitive way. The first ARPANET link was established between the University of California, Los Angeles, and the Stanford Research Institute on October 29, 1969, and the Specification of Internet Transmission Control Program (RFC-675) [22] by Vinton Cerf, Yogen Dalal and Carl Sunshine, in December 1974, contains the first attested use of the term Internet. The current Internet Protocol, IPv4 was described in RFC-791 on September 1981 [165]. Since then, Internet has grown exponentially, very slowly until the 90s when it began to take off (Figure 1).

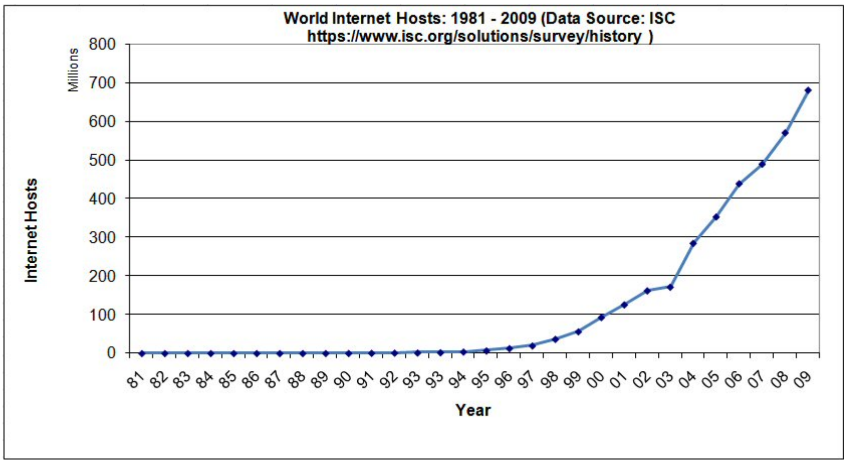


Figure 1: Internet growth from 1981 to 2009

Unfortunately the poor assignation policy of IPv4 addresses for several years has led to IPv4 address exhaustion which will probably take place in the second semester of 2011, even though some sources say it will be even earlier. In 1998 the Clinton administration created the ICANN [88] to regulate this issue and to preserve the operational stability of the Internet, but the damage had already been done. The new generation Internet protocol, IPv6, is broadly accepted as the definitive solution for the IPv4 shortcomings, including address exhaustion. Even though IPv4 address exhaustion was detected long ago and now it is really closing in, transition towards IPv6 has been really slow for many years. Only in the last few years, since 2004, have many stakeholders promoted and developed IPv6.

Furthermore, Cisco System has predicted an Internet bandwidth annual growth of 42% and the word exaflood is beginning to make sense. This growth is due to new user trends which include mobility (WiMAX, HSDPA and LTE in the horizon), IPTV and IT convergence. The fact that YouTube currently consumes as much bandwidth as the whole Internet did in year 2000, is worth emphasizing. There has actually been a convergence of transport platforms (e.g. broadcasting networks and Internet) and a convergence of terminals (e.g. television set and PC). Through this convergence broadband has 'de facto' already become an essential digital utility for many households and businesses. A new service convergence architecture, the IP Multimedia

Subsystem (IMS), was developed by the 3rd Generation Partnership Project (3GPP) and has been accepted and integrated in the new architectures developed by different organizations. This new architecture, IMS, is empowered by content and application providers as long as operators because of the cost's optimization.

User expectations of Internet have evolved with time and nowadays users are requesting Quality of Experience (QoE) rather than simply a connection. This could be achieved by raising significantly the available bandwidth per user, but this it is not that simple neither at the user end nor at the core. Moreover, if the battle for net neutrality is won by service providers this QoE might be suffer extreme consequences.

Current access networks have become traffic bottlenecks. Therefore the deployments of NGAs are increasingly becoming a reality in order to respond to user demands of bandwidth for the new on-line services and applications. NGAs provide not only high speed Internet but a quality physical connection, increased consistency and reliability, and decreased latency and bit error rate. In the medium-term, the NGAs will be the essential infrastructure for national competitiveness and will form part of the standard of living of all the citizens of a country. Unfortunately for service providers, there is a rebound effect of adopting broadband and/or increasing bandwidth rates, given that according to a study by Nielsen/Netrating Netview, greater capacity leads to greater usage. At present, bandwidth is very rarely guaranteed by NGA operator offers and only some Asian operators are beginning to integrate QoS into their offers (e.g. Japanese NTT). The exponential growth of the Internet has made that in the short-term the bottlenecks will also begin to appear in the core besides the access [171].

Therefore, QoS is more necessary now than ever before. The central problem in proving consistent end-to-end QoS services is the difficulty in configuring network devices like routers and switches to handle packet flows in a manner that satisfies their requested QoS requeriments. This problem is especially acute when the end-to-end data path of an IP QoS session crosses multiple administrative domains managed by different operators because every single network along the end-to-end data path must provide the appropriate QoS regardless of the technology of each network. Although operators may agree on the QoS requeriments of a particular set op IP services, they may not configure their network devices in the same way to implement the services due to differences in their network topologies, QoS mechanisms available in the network devices, and other non technical/management requeriments. Therefore, there is a need to create solutions that permit network operators to easily configure their networks to implement consistent QoS services without dealing with the complexity of their networks.

Furthermore, this QoS service must be provided on a seamless end-to-end basis, with special attention when multimedia flows traverse heterogeneous networks belonging to different operators. Thus, a QoS mapping will have to be applied in the border nodes of each domain in order to assure this seamless QoS and a method of coordinating resource allocation between one autonomous system and another will also be required.

All these requirements have been included in the development of the new architectures of Internet, the Next Generation Networks. This NGN concept takes into consideration the new realities in the telecommunication industry characterised by factors such as the need to converge and optimise the operating networks and the extraordinary expansion of digital traffic.

1.1.1. Networking research enviroment

This thesis has been submitted to the PhD program Tecnologies de la Informació i les

Comunicacions i la seva gestió (Doctorate in Information and Communication Technologies and their Management) of the Universitat Ramon Llull (URL). This thesis is part of the research of the Grup de Recerca en Telemàtica (Research Group in Networking Technologies), of Enginyeria La Salle of the URL.

The Research Group in Networking is mainly focused on techniques related to QoS and routing. Since its beginning in 2004 it was decided to focus group activity on the WLAN and PLC access networks as well as on the core networks with MPLS, IPv6 and QoS. Moreover the departmental proximity of the Grup de Recerca en Sistemes Intel·ligents (GRSI) research group (Research Group in Intelligent Systems) has oriented part of the activity towards QoS optimization using artificial intelligence techniques which is the main topic of research of the GRSI. In this field genetic algorithms have been applied in the Networking research group to optimize traffic engineering in routing protocols. The Research Group in Networking seeks to apply its research in real world domains which may benefit society in the near future. Currently, the group is working on routing, mesh networking and QoS over PLC and the Smart Grid.

The research work in the QoS domain of this thesis' author started in 2004 with the project “Telefonia sobre IP con calidad de servicio sobre accesos xDSL, PLC y WLAN” (Telephony over IP with Quality of Service over xDSL, BPL and WLAN accesses), partially funded by the Ministerio de Industria, Turismo y Comercio under grant number FIT-330200-2004-189. In this project the key performance QoS indicators (bandwidth, delay, jitter and packet loss) were evaluated over a ToIP testbed. This was the first approach to the QoS, the WLAN and PLC access networks, the SIP protocol and the research from a practical point of view. The state of the art of the IPv6 protocol analysis presented in this thesis was also developed in this project.

This work continued with the project “CSI-RHET: Calidad de Servicio e Ingeniería del Tráfico en Redes Heterogéneas” (Quality of Service and Traffic Engineering in Heterogeneous Networks) in 2006 funded by the Ministerio de Educación y Ciencia under grant TEC2005-08068-C04-04/TCM of the Plan Nacional I+D+i. The line of research was focused on addressing the design, implementation and evaluation through simulation and experimental testbeds of a heterogeneous multiservice broadband telecommunications networks based on IPv6 with QoS and Security concerns, where the paradigm of architecture and Internet service provisioning is inherent and interoperability with current network access technologies is supported. The QoS testbed over IPv6 networks and the QoS management architecture presented in this thesis were developed in this project. This project had four partners: Universidad de Cartagena, Instituto Tecnológico de Aragón, Universidad de Alcalá de Henares and Universitat Ramon Llull.

The line of research was later continued within project OPERA 2: Open PLC European Research Alliance for New Generation PLC Integrated Network Phase 2 beginning in 2007 and funded by the European Commission's Sixth EU Framework Programme for Research and Technological Development (FP6) under grant FP6-IST-OPERA (number 026920). This project aimed to develop PLC technology as a real alternative to existing broadband technologies and to enhance the safety of communications and supply a real alternative to the customers while increasing competitiveness. The QoS management architecture for PLC access networks presented in this thesis was developed in this project. This project had 26 partners.

A visual review of this thesis's author research evolution since 2004 can be seen in Figure 2.

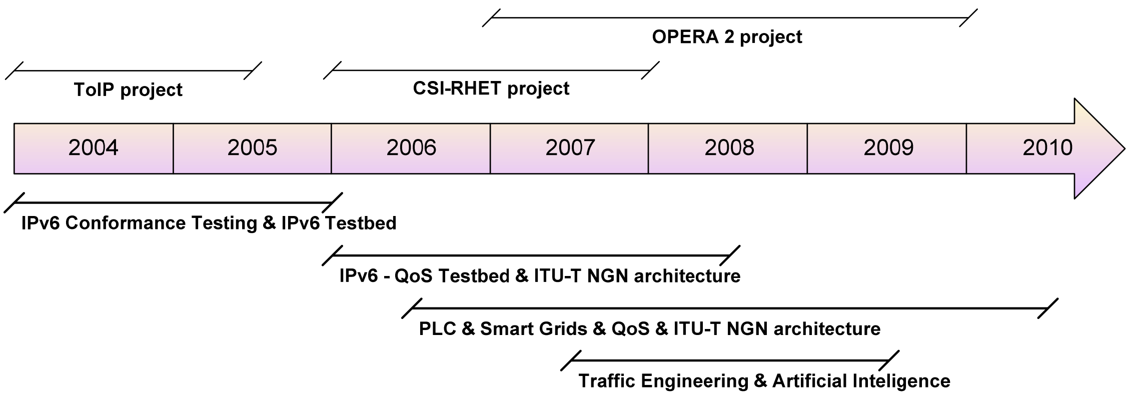

OPERA 2 project
ToIP project
CSI-RHET project
2004
2005
2006
2007
2008
2009
2010
IPv6 Conformance Testing & IPv6 Testbed
IPv6 - QoS Testbed & ITU-T NGN architecture
PLC & Smart Grids & QoS & ITU-T NGN architecture
Traffic Engineering & Artificial Inteligence

Figure 2: Timeline diagram of the research lines and projects

The Research Group in Networking has proposed several other projects, some of which have been accepted, like the European FP7 INTEGRIS project that began in February 2010, and some of which have not, like the ToIP2 PROFIT project or the IMMOTEQ and ISOM projects from the *Plan Nacional I+D+i*. This latter was focused on developing a Path Computation Element (PCE) device for managing DiffServ aware MPLS Traffic Engineering (MPLS DS-TE) networks and applying artificial intelligence techniques for the path computation with QoS over a MPLS DS-TE experimental testbed with open source software. The group has also participated in the three phases of the ENIGMA project [45], related with MPLS-TE technology and co-participated by I2CAT, VODAFONE, Universitat Politècnica de Catalunya and Universitat Ramon Llull. Part of this thesis develops the initial steps of this approach, including the analysis and practical implementations of applying artificial intelligence techniques to traffic engineering in NGN networks. Further steps, including the state of the art of the MPLS DS-TE implementations, with preliminary results are presented in the future work and the appendixes.

1.2. Thesis objectives

The general goal of the present work is to provide practical contributions to the data networks QoS field and to outline the advantages of intelligent systems. As stated in previous sections, there is a growing need for Quality of Service control architectures in order to provide consistent and seamless end-to-end QoS services in the environment of the Next Generation Networks. Furthermore, these architectures must be so generic that they can be adapted to any access network (especially NGAs), as well as core networks and they must also support the IPv6 protocol which is mandatory for NGN architectures.

This thesis proposes a practical view of the research and is therefore focused on implementing rather than simulating, even though simulations are provided when necessary. In this line of work, GNU/Linux solutions are used on the research path due to their flexibility. In fact, the research lines set out in this thesis have been defined mainly by the objectives of the research projects developed in the aforementioned research activity which adhere to the guidelines of the Research Group in Networking Technologies.

Specifically, this leads to the definition of the following four objectives:

1. Perform conformance tests on exiting IPv6 protocol implementations and implement a testbed supporting IPv6.
2. Design and implement an automated QoS management architecture for NGNs.

3. Propose a QoS control architecture for next generation Broadband Power Line Networks.

4. Study and implement traffic engineering with artificial intelligence in NGNs

As follows, each one of the four objectives is elaborated in detail.

Perform conformance tests on exiting IPv6 protocol implementations and implement a testbed supporting IPv6. In order to build a testbed using the IPv6 protocol it is fundamental to assure the correct performance of the devices included in the testbed. When this thesis started in year 2004 there were serious doubts about the correctness of the IPv6 implementations on the market. The first objective of the thesis is to implement a testbed supporting IPv6 and therefore a thorough analysis of the state of the art of the IPv6 conformance and the interoperability testing needed to provide performance guarantees according to specifications must be conducted. The empirical study by conformance tests of the IPv6 protocol implementations in end node systems must be carried out if they want to be used as routers in the IPv6 testbed.

Design and implement an automated QoS management architecture for NGNs. The second objective is to develop a QoS broker architecture, with a QoS broker device to manage QoS services and a tested with NGN features in order to empirically demonstrate the proper functioning of the system. For this purpose, we initially propose the development of a bandwidth broker architecture, which performs automated DiffServ (Differentiated Services) management to the DiffServ nodes located in a QoS testbed, enhanced from the IPv6 testbed for QoS support, and it allows intradomain and interdomain communication therefore providing end-to-end QoS. The second step is to update the initial architecture to fully support the NGN proposed specifications by the ITU-T for an end-to-end QoS control architecture.

Propose a QoS control architecture for next generation Broadband Power Line Networks. The ITU's NGN architecture takes into account the access domains (NGAs), but not all of these domains are ready to manage QoS under the necessary end-to-end requirements. The broadband PLC is a key access technology which is about to be standardized and the OPERA project specification has opted to be the new IEEE P1901 standard for PLC. The third objective of this thesis is to develop a proposal for the integration of the QoS management in the new broadband PLC access networks proposals within the ITU-T architecture. This proposal must meet the OPERA project specifications but be generic enough to be applied to the other broadband PLC access networks proposals. Moreover, the QoS control necessities in the Smart Grids' communications network architecture have also to be analysed, even though a detailed solution cannot be proposed as it is still at its first stages of definition.

Study and implement traffic engineering with artificial intelligence in NGNs. In the last objective of this thesis, we provide a solution to optimize routing in NGNs with artificial intelligence. The use of these techniques is empowered by the proximity of the GRSI research group and these techniques have been proposed to enhance the overall QoS in most of the research projects we have participated in. These algorithms provide certain autonomous intelligence to the QoS Broker and therefore have to be implemented in the QoS Broker in order to optimize traffic engineering and/or LSPs in NGNs.

Each one of these objectives is dealt with in a separate chapter of the present thesis. The overall structure of the document is provided in the following section.

1.3. Road map

This thesis is organized, in addition to the present chapter, in six further chapters. A brief description of the chapters included in this thesis is detailed below:

Chapter 2 provides an in-depth review of the state of the art of the IPv6 protocol and its stakeholders. It also provides a deep study of the worldwide IPv6 conformance testing initiatives and the scope of the IPv6 Ready Logo program certifications. An IPv6 conformance test study is carried out over IPv6 implementations in end node operative systems and the results are provided, in both an analytical form and in a practical form, from the IPv6 testbed implemented with GNU/Linux computers.

Chapter 3 focuses on the QoS policy-based network management. A broad overview of the QoS technologies is first provided and then the automation of QoS management is analysed and the protocols that suit the requirements are discussed. The design and implementation of a bandwidth broker device for managing Diffserv domains with IPv6 networks is presented next. The design has required the development of a new protocol named SIBBv6 for bandwidth broker signalling between peers. Finally the evaluation results of applying this bandwidth broker over the enhanced IPv6 testbed with QoS capabilities are provided.

Chapter 4 turns to the facetwise analysis to the Next Generation Network architecture proposed by the ITU-T. This architecture is still under development and, thus, protocol discussion is provided on the basis of the early requirements by ITU-T. A QoS broker device to manage QoS seamless end-to-end QoS through heterogeneous domains is proposed. The evaluation methodology has cloned the experiments carried out in chapter 3 over the same testbed but using different protocols according to the ITU-T requirements.

Chapter 5 proposes applying the QoS broker architecture for NGNs into the BPL access networks. The new generation of PLC architectures is still under development and, therefore, a theoretical proposal for the BPL architectures, proposed by the OPERA project and the IEEE P1901 working group, are discussed. This chapter also overviews the Smart Grids and the possible applications of the proposed QoS broker architecture over this future technology.

Chapter 6 presents a proposal for optimizing traffic engineering using artificial intelligence. The state of the art of the genetic algorithm used for traffic engineering is presented and a simulation study is carried out in order to determine the best algorithm. Finally, the details of the experiments carried out with a QoS broker with an hybrid genetic algorithm implemented to evaluate enhance the performance in real data networks using commercial Cisco routers are provided. The testing scenarios and experimentation results are also detailed.

Chapter 7 finishes with the contributions of this thesis by summarizing, providing key conclusions, reviewing the main lessons extracted from this work, and by suggesting future work lines.

2. IPV6 CONFORMANCE TESTING

2.1. Introduction

The current version of the Internet Protocol, IPv4, has been almost unaltered since its publication in 1981 in RFC-791 [165]. For more than 20 years IPv4 have proved to be robust and scalable. However, the exponential growth of the Internet in volume as well as in complexity has taken de IPv4 design to its limits. In 1992, the first symptoms of IPv4 address exhaustion started to be evident. The progressive depletion of the IPv4 addressing space, the complexity of multi-homing and the user requirements for mobility, security and QoS are the mid term unaffordable challenges for the actual Internet Protocol version 4. The next generation Internet Protocol, IPv6, is widely accepted as the definitive solution to the weaknesses of the current IP protocol.

The shortage of IPv4 addresses is especially acute in Asia, where the rapid growth of Internet connectivity cannot be served by the IPv4 address space that has been assigned. In 2004 almost 74% of the addresses had been assigned to North American organizations, with two universities (Stanford and MIT) having more addresses assigned to each one of them than the Peoples' Republic of China, with 80 million Internet users, for the only reason that the Internet revolution started later in Asia. Therefore, this region is particularly interested in the development of IPv6 and some Asian governments have even announced their official support of this technology and have supported their national IPv6 Promotion Councils. The Japanese government took the political lead in the region in March 2001 with the publication of the e-Japan Priority Policy Program which set the expired deadline on the end of 2005 for the upgrading to IPv6 of every existing network in business and public sectors. South Korea and Taiwan followed the Japanese initiative announcing plans to roll out IPv6, and bilateral consultations, at ministerial level, between P.R. of China and Japan have taken place by means of further IPv6 promotion. In 2010 China has more than 400 million Internet users with a 30% inteannual growth [101].

Even though it seems unstoppable, protocol migration has been so far surprisingly slow. The appearance of killer applications would cause an acceleration of this process, but experts have yet to conclude whether they are non-existent or simply still have not taken off in the market. Ubiquity with Mobile IPv6 (MIPv6), multimedia communications with end-to-end services or the new IMS services could be some examples of such applications. It is clear, though, that With IPv6 there will not be an immediate transition as it happened in 1981 with IPv4, when the Network Control Protocol just disappeared. The vast scope reached by the IPv4 Internet will make it impossible. The protocol migration, due to the never ending postponement of the IPv4 address depletion, will probably be slow but unstoppable during the next years. Currently, Internet is a dual system that may work with IPv4 or IPv6, even though IPv6 is still residual (Figure 3).

Figure 3: The IPv4 to IPv6 migration’s evolution

One possible reason for the slow pace of the IPv6 deployment is that the IPv6 Internet requires the upgrading of many important network elements such as node operating systems, network elements, management systems and ISPs, carrier operating procedures (security, troubleshooting, etc.) as well as the necessary migration of all the applications with network modules. Another important aspect to consider is that the aforementioned elements have to demonstrate interoperability, something that cannot be achieved without conformance testing.

Probably the most important fact at present time is the almost completion of the global standardization of the IPv6 specifications by the IETF’s working groups [93] related with the IPv6 protocol. Most of them have concluded their tasks and closed or are in the last stages. The ultimate aim of standardization is to enable the interoperability of different implementations and this interoperability can not be reached without conformance testing of the implementations according to the specifications. The two worldwide initiatives related to IPv6 conformance testing have completed most of their tasks. The IPv6 Ready Logo program [104] is about to approve the phase 3 logo and many vendors have already obtained the phase 2 logo. The ETSI IPv6 Testing Project [58] has closed phase 1. The first chapter of this thesis focuses on the interoperability and conformance testing issues related to IPv6.

2.1.1. Objectives

The fundamental objective of this first line of research was to initiate the research path by opening a new research line in the latest Internet technologies. Working from this basic premise, the first objective of the scientific research was both, formal and methodological, about the scientific research. It was to achieve a good research methodology and to develop the capacity to write scientific papers of enough quality to be accepted by the international scientific community.

The second objective aimed to provide a deep knowledge of the IPv6 protocol and its stage of development. This consisted of compiling information on the level of formal maturity of IPv6 from the perspective of state of the art of the standardization actors, promoters, developers and evaluators of the IPv6 technology, as well as the interoperability and conformance testing issues related to IPv6.

The last objective of this phase was to achieve practical knowledge. The aim was to carry out IPv6 conformance tests of different operative systems which support IPv6 which would lead to the election of one of them to be implemented in an IPv6 testbed. Therefore, the final realization had to be a stable IPv6 testbed with performance guarantees from which to open new research lines could be opened.

The success of this initial phase is measured by the quantity and quality of the objectives achieved which include the acceptance of at least an international scientific paper and the implementation of an operative IPv6 testbed.

2.2. State of the Art of the IPv6 protocol

Internet standards are established by a group of organizations, all of which operate under the auspices of the Internet Society (ISOC) [110]. ISOC is the organizational home of the Internet Engineering Task Force (IETF), the Internet Architecture Board (IAB), the Internet Engineering Steering Group (IESG) and the Internet Research Task Force (IRTF). At a technical and developmental level, the Internet is made possible through the creation, testing and implementation of Internet Standards. These standards are developed by the IETF working groups and then are considered by the IESG, who is responsible for the technical management of IETF activities and the Internet Standards Process, in consultation with the IAB. The RFC Editor, supported by the Internet Society, is responsible for preparing and organizing the standards in their final form.

The IETF [93] is a global non profit organization encouraged to develop the underlying technology in the Internet. As stated in the RFC-3935 [2], its goal is to make the Internet work better by producing high quality and relevant technical and engineering documents that influence the way people design, use and manage the Internet. The IETF is a global non profit organization with sole authority over the IPv6 protocol standards.

During the development of a specification by a working group, draft versions of the documents are made available for informal review and comment by placing them in the IETF's Internet-Drafts directory. A draft specification has to be placed in the RFC Editor’s Queue for its approval and final publication to become a RFC. Before a specification can become a Proposed Standard it needs to be published as part of the Request for Comments (RFC) document series which is the official channel for many publications of the Internet community. This RFC Editor is responsible of its final format. Once a specification has been published as a RFC, then it may become an Internet Standard if it follows the “Internet Standards Process” described in RFC-2026 [16]. This process is currently used by the Internet community for the standardization of protocols and procedures. This process has three Standard Track levels. A Proposed Standard is a stable specification that is believed to be well-understood and should have no known technical omissions in relation to the requirements. A Draft Standard is a specification which has been implemented independently at least by two different sources which have to be completely interoperable. Finally, a Standard, labeled as a STD, is the specification for which significant implementation by different stakeholder and successful operational experience has been obtained.

The IETF working groups are organized into several areas according to topics. Many of these working groups consider IPv6, but the main one is the IPv6 Working Group [96].

2.2.1. The IPv6 specification

The IPv6 protocol was developed by the IETF as a response to the obvious lacks of the IPv4 protocol. Actually, IPv6 is a group of protocols and standards, which are developed around the core specification.

At the beginning of 1990, the IETF started studying the problem of expanding the Internet addresses and in 1994 released a first official recommendation for a new generation Internet protocol maned IPng (Internet Protocol next generation).

In January 1995, the Recommendation for the IP Next Generation Protocol (RFC-1752) [15] was published. This specification describes the requirements for this new protocol, IPng, and specifies the PDU format and the addressing procedures, the routing and the security.

In December 1995, the Internet protocol version 6 specification (RFC-1883) [36] was published. This specification became obsolete in 1998 with the publication of the RFC-2460 [37], currently a Draft Standard. This is the protocol’s general specification but there are many more which specify the concrete details, like the one specifying addressing [82].

2.2.2. IPv6 working group at IETF

The IETF is a non-profit global organization with sole authority over the IPv6 protocol standards. These standards are developed by the IETF working groups (WG), mainly by the IPv6 WG, though most WGs within the IETF have contributed to the IPv6 progress in their technological area. Originally chartered as the Internet Protocol Next Generation working group (IPng), this group is currently known as the IPv6 WG. It began implementing its recommendations after the IETF meeting of July 1994 and the publication of “The Recommendation for the IP Next Generation Protocol” (RFC-1752) in January 1995. The aim of this WG is the standardization of IPv6 and other related protocols. The main issues concerning this WG are addressing, management and control. All the documents edited by this WG which could be standardized are already in the Standards Track Process (as Proposed or as Draft Standards) or have been released as Informational or Experimental documents. The most important is the Draft Standard specification of the Internet Protocol version 6 (RFC-2460).

By December 2006, despite the fact that the last issued drafts had been submitted, some of which had already been accepted in the RFC Editor’s Queue (many for updating former RFCs), two of them still remained active and unclosed. These active drafts revolved around some text editing within RFC-2461bis and RFC-2462bis in relation to DHCP address configurations in the

IPv6 Router Advertisement. By that moment in time almost all the goals and milestones had been reached, except for the last one, consisting on the re-chartering of the WG and its closing, even though it was originally planned for it to be closed by November 2005. In words of the WG Chair, Brian Haberman, “the group is extremely close to achieving all its goals”.

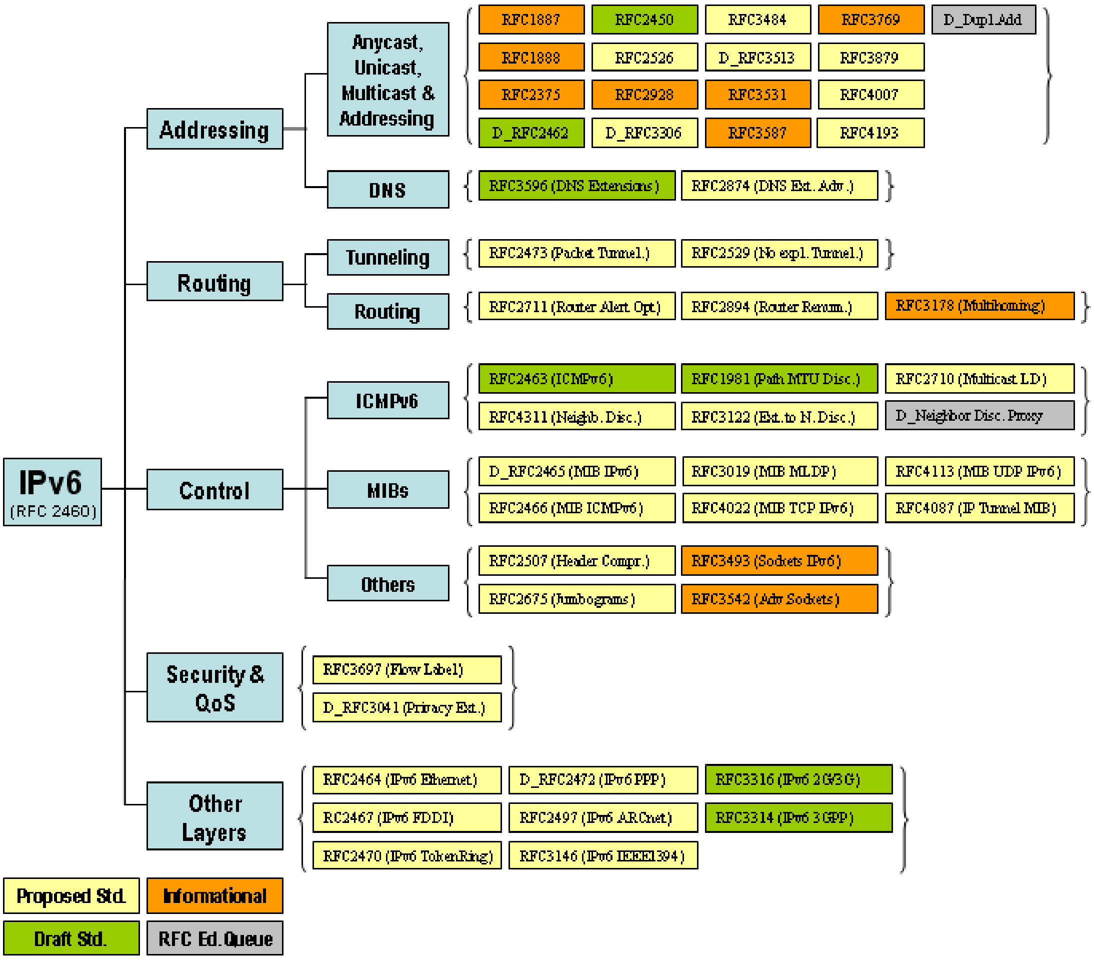


Figure 4: The IPv6 Working Group specifications in December 2006

Figure 4 shows all the specifications developed by the IPv6 WG, the distribution of RFCs and drafts according to the fields they belong to and the status of each document by December 2006. It is organized by colors showing the status of the document: Draft Standards are in green, Proposed Standards in yellow, Informational RFCs in orange and, finally, Experimental RFCs in violet. Some RFC appear with a preceding “D” which stands for a RFC made obsolete by an existing draft.

There are many other working groups in the IETF with IPv6 related milestones. Table 1 shows the status of the most important among them from the relevance point of view. Most of them are already closed or are scheduled to close soon. It must be noticed that whilst active drafts waiting to become RFCs exist the group is considered to be active.

Some of these working groups have resumed or updated the task of concluded working groups. For example the IPv6 WG has updated some specifications mainly related to MIBs.

WG	Field	Scheduled conclusion	Status	RFCs	*Standards Track docs*
MIPSHOP	Mobile IPv6	September 2005	Active (delayed)	✓	-
MIP6		August 2006	Active (delayed)	✓	✓
MONAMI6		June 2007	Active (delayed)	-	-
MULTI6	Multihoming	December 2004	Active (delayed)	✓	-
SHIM6		June 2006	Active (delayed)	-	-
6LOWPAN	WPAN	July 2005	Active (delayed)	✓	-
NGTRANS	Transition	February 2003	**Concluded**	✓	✓
V6OPS		July 2005	Active (delayed)	✓	-
RIP	Routing	April 2002	**Concluded**	✓	✓
OSPF		November 2003	Active (delayed)	✓	✓
VRRP		November 2004	Active (delayed)	✓	✓
IS-IS		November 2005	Active (delayed)	✓	✓
BGP (IDR)		October 2006	Active (delayed)	✓	✓

Table 1: Other IETF IPv6 related working groups in December 2006

2.3. State of the art of the stakeholders

There are many organizations and consortiums dedicated to IPv6 promotion. The main worldwide consortium is the IPv6 Forum [103], a non-profit organization formed by the leading Internet vendors and National Research and Education Networks (NRENs). Its mission is to promote IPv6 by improving market and user awareness of this technology. The IPv6 Ready Logo Program, one of the several initiatives launched by the IPv6 Forum, will be discussed later in this thesis.

Many regional organizations also actively promote IPv6, such as the IPv6 Task Forces (IPv6TF) and the Asian IPv6 Promotion Councils. The most important IPv6TF are the North American IPv6 Task Force (NAv6TF) [150], a sub-charter of the IPv6 Forum, and the European IPv6TF [51] led by the European Commission. The Japanese IPv6 Promotion Council is the leading representative of the Asian Promotion Councils and it is also involved in the development of some of the newest IPv6 conformance test specifications.

The most important working project to develop IPv6 is the WIDE project [216], a research consortium between industry, public institutes and academia in Japan founded in 1985. WIDE focuses on global Internet operations and is at the forefront of pioneering IPv6 practical development projects, such as KAME [127], USAGI [200] or TAHI [188]. These projects closely collaborate to offer free open-source IPv6 codes and tools for different platforms. Their contributions include "the operation of the M Root DNS Server, forming the Internet backbone in Asia-Pacific region and the world's first undertaking of IPv6" [216].

The TAHI project started in 1998 and is a joint program led by the University of Tokyo and the Yokogawa Electric Corp. with the objective of developing and providing open source conformance and interoperability tools and tests for IPv6. Basically it enhances efficiency in the development of this technology. This project provides the most advanced IP verification packages for free and is one of the bases of the IPv6 Ready Logo Program.

Table 2 summarizes the URLs of these projects and the principal IPv6 testing related sites in the Internet, all of which are mentioned in this chapter.

6Bone	www.6bone.net	IRISA	www.irisa.fr
Agilent Technologies	www.agilent.com	IXIA	www.ixiacom.com
CNGI project	www.edu.cn	Japanese IPv6 Promotion Council	www.v6pc.jp
eEurope Action Plan	www.e-europestandards.org	Microsoft Corporation	www.microsoft.com
ETSI	www.etsi.org	Moonv6 project	www.moonv6.org
European IPv6TF	www.ipv6tf.org	NAv6TF	www.nav6tf.org
GÉANT2 project	www.geant2.net	Net-O2 Technologies	www.net-o2.com
IETF	www.ietf.org	Spirent Communications	www.spirentcom.com
IPv6 Forum	www.ipv6forum.org	UNH-IOL	www.iol.unh.edu
IPv6 Ready Logo program	www.ipv6ready.org	WIDE, TAHI, KAME i USAGI projects	www.wide.ad.jp
IPv6 Testing project	www.ipt.etsi.org		

Table 2: URL of the main stakeholders related to IPv6

2.3.1. End node Operative Systems

In order for IPv6 to succeed, there is an underlying need for IPv6-enabled end node equipment. Most of the widely used operative systems, such as the Apple MAC OS X which has incorporated IPv6 supports since version 10.2, have introduced IPv6 supports in their latest releases. The main contributors to IPv6 development in this field have been WIDE and Microsoft Corp.

One of the first end node OS to support IPv6 features was UNIX. The KAME project was launched in 1998 and provided the most advanced IPv6 code in a free and operative IPv6 and IPSec stack for UNIX BSD variants: FreeBSD (starting from version 4.0), OpenBSD (since version 2.7), NetBSD (since version 1.5), BSD/OS (since version 2.4). These distributions already have KAME code merged in the kernel. The KAME project concluded in March 2006 with the successful completion of its project mission.

The Linux OS started to offer experimental IPv6 support in 1996. The USAGI project (UniverSAl playGround for IPv6) was launched in 2002 and its aim is the development of updated implementations of IPv6 and IPSec for the Linux OS. USAGI publishes two stable kernel releases per year with the most advanced IPv6 implementations as well as unstable snapshots every two weeks. This code is currently integrated into the Linux official kernel (since version 2.6).

Microsoft Corp. [140] has evolved its IPv6 stack since the first release for Windows NT in 1998 (named MSRIPv6). The IPv6 Technology Preview was published in 2000 for Windows 2000 and the IPv6 Developer Preview in 2001 for Windows XP. The IPv6 stack is already fully integrated into the latest versions of Windows OS.

Apple's MAC OS X currently supports IPv6 (since version 10.2), but this UNIX based OS was one of the latest to be qualified for IPv6 Ready Logo Program certification.

2.3.2. IPv6 networks

The main IPv6 networks are: 6Bone which was the first global IPv6 Internet, MOONv6 in the U.S.A., GÉANT2 in Europe and CERNET2 in China. However there are many other IPv6 networks, such as the Japan Gigabit Network.

The 6Bone [66] was an IPv6 Testbed network launched in 1996 by the IETF to allow IPv6 testing and IPv6 transitioning into the Internet that interconnected NRENs from several countries as well as commercial companies involved in the development and testing of IPv6. As a result of agreements with the IETF IPv6 community, the assigned 3FFE test prefixes can not be used after June 2006 due to the existing large IPv6 production deployment. Therefore, it is has been phased out and all the test prefixes have been returned to the ICANN and all the infrastructures that remain from the old 6Bone have been moved to production address space.

The MOONv6 project [199] started in 2003 and is an initiative led by the NAv6TF, involving the InterOperability Laboratory at the University of New Hampshire (UNH-IOL) [198], the Internet2 project (I2) [102] and the U.S. Department of Defense. It is a multi-site native IPv6 network designed for testing the interoperability of the real network requirements of various vendor-specific IPv6 implementations.

The GÉANT2 project [38], co-funded by the European Commission and Europe's NRENs, began in 2004 and is the next generation European research network. The GÉANT2 network connects 34 countries through 30 NRENs since 2008. Even though GÉANT2 initially operates a dual-stack IPv4/IPv6 service, the majority of NRENs connected to GÉANT have IPv6 as a native protocol in the network core.

The Chinese Next-Generation Internet demonstration project (CNGI) [26] was created by the Chinese government and started operating in 2003. Its main achievement is CERNET2, the first IPv6 backbone network in China, which is the only nationwide academic network in China and interconnects 25 universities in 20 cities. It is expected to support 100 universities in the near future.

2.4. Conformance and interoperatibity testing

Conformance testing is the verifying process for an implementation according to the specification on which it is based. The full IPv6 support involves an important amount of specifications, most of which are already standardized or at the final stages of standardization. Although many studies have discussed protocols conformance testing in the past, nowadays most of the development is based on the ISO/IEC 9646 conformance testing methodology [109][131]. Some freely available and commercial IPv6 conformance testing suites exist, which will be addressed later in this chapter.

On the other hand, interoperability testing is the process of checking whether two implementations are able to work together effectively, without prior communication, in order to exchange information in a useful and meaningful manner. The first step towards full interoperability is conformance. Once an implementation has proved to be compliant with the IPv6 specifications, it should be tested in a multi-vendor environment. This can be done in specialized laboratories or at the various interoperability events which take place throughout the year. It is worth mentioning that these events are critical to the advancement of the standards themselves since they provide a necessary feedback to standards developers.

There are two worldwide initiatives related to IPv6 testing: the IPv6 Ready Logo Program and the IPv6 Testing Project.

2.4.1. IPv6 Ready logo program

With the aim of accelerating the development and homologation of IPv6 capable devices, the IPv6 Forum launched the IPv6 Ready Logo Program in 2003. The importance of the program lies in the generally accepted idea that a globally unique logo program should be defined in order to avoid confusion in customers' minds. Thus contributing to the perception that IPv6 is available and ready to be used with the IPv6 Ready logo. The increasing number of IPv6-enabled implementations over the last few years has forced remaining hardware/software vendors to react and to improve their IPv6-related products. As interoperability has always been considered as a critical feature in the Internet community, it is very important to provide the market with a strong signal to demonstrate the level of interoperability between products from different vendors. The other objective of this program is to create the capacity to guarantee minimum levels of interoperability.

Figure 5: Silver and Gold logo of the IPv6 Ready Logo program

To become certified with the IPv6 Ready logo, an IPv6 implementation must show full compliance with the IPv6 specifications and must also prove interoperability with other IPv6 implementations. When the product does not satisfy the test specifications, the product can not be certified, even if it has highly advanced functions and interoperability with many other products. The logos can be delivered to different product categories from final products to stacks in either phase. The devices may pass the certifications as a host, as a router or as a special device (e.g. IP camera).

Program phases

The IPv6 Ready Logo Program plans IPv6 full compliance in three phases, which are related to three IPv6 Ready logos and are obtained depending on the grade of compliance and requirements stated by the IPv6 Program Committee (v6PC).

The first phase started in September 2003, with the TAHI project test package, and its goal is for all its implementations to support basic IPv6 features with minimum guarantees. There are two main requirements within the criteria in order to obtain the phase 1 (silver) logo. The first requiriment forces the device being tested (node under test or NUT) to support 100% of IPv6 mandatory core protocols (RFC-2460, 4861, 4862 and 2463) of the test specifications to pass the self-test. These RFCs are related to:

- IPv6 Specification (RFC-2460) [37]
- Neighbor Discovery (RFC-4861) [149]
- IPv6 Stateless Autoconfiguration (RFC-4862) [194]
- ICMPv6 (RFC-4443) [32]

The second requiriment of the methodology forces the NUT to prove being interoperable with 4 other IPv6 devices from other vendors including 2 hosts and 2 routers to pass the interoperability test. Figure 6 shows an example of interoperability test network with some Test Nodes (TN) and Test Routers (TR).

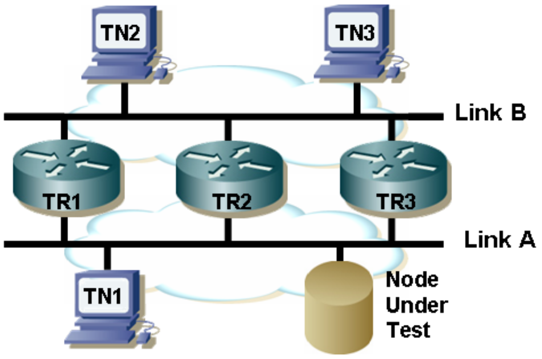
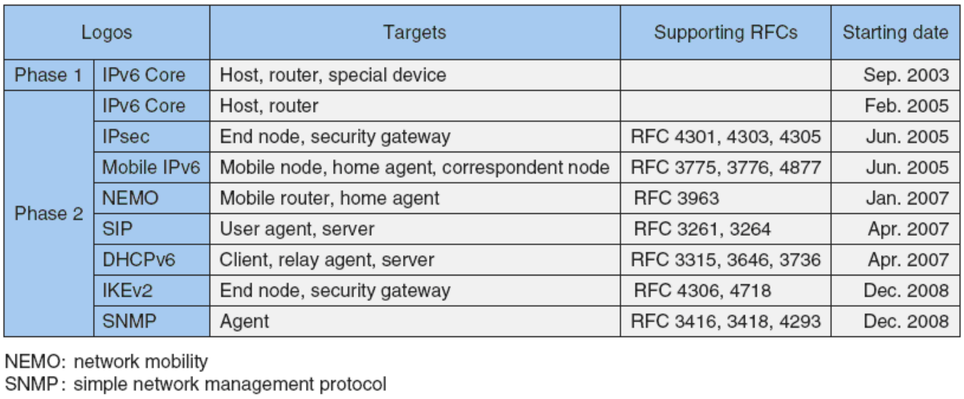


Figure 6: Interoperability scenario

The second phase started in February 2005 and its goal is to certify the full IPv6 core protocols support and some optional advanced features. The requirements needed to pass the tests and to obtain the phase 2 logo (golden logo) are the same as in phase 1 but with some more accurate conformance and interoperability tests and specifications, such as RFC-1981 related to PMTUD (Path MTU Discovery) [137], IPSec (Internet Protocol Security), MIPv6 (Mobile IPv6), SIP (Session Initiation Protocol) and DHCPv6 (Dynamic Host Control Protocol for IPv6).

The IPv6 Ready Logo program phase 2 is generic in order to avoid confusion and is given to those which pass the mandatory IPv6 core protocols tests that include the complete set of IPv6 core specifications (RFC-2460, 4861, 4862, 4443 and 1981).

There are six additional phase 15 optional logos including IPSec (2), MIPv6 (3), NEMO (2), SIP (2), DHCPv6 (3), IKEv2 (2) and SNMP (1) as can be seen in Figure 7.



Logos		Targets	Supporting RFCs	Starting date
Phase 1	IPv6 Core	Host, router, special device		Sep. 2003
Phase 2	IPv6 Core	Host, router		Feb. 2005
	IPsec	End node, security gateway	RFC 4301, 4303, 4305	Jun. 2005
	Mobile IPv6	Mobile node, home agent, correspondent node	RFC 3775, 3776, 4877	Jun. 2005
	NEMO	Mobile router, home agent	RFC 3963	Jan. 2007
	SIP	User agent, server	RFC 3261, 3264	Apr. 2007
	DHCPv6	Client, relay agent, server	RFC 3315, 3646, 3736	Apr. 2007
	IKEv2	End node, security gateway	RFC 4306, 4718	Dec. 2008
	SNMP	Agent	RFC 3416, 3418, 4293	Dec. 2008

NEMO: network mobility
SNMP: simple network management protocol

Figure 7: IPv6 Ready logos

Each one of these technologies has defined its own test profiles with associated requirements. Tests on MLD (Multicast Listener Discovery) and transition (6to4 and NAT-PT) technologies are under development. They will also be considered as optional with their specific phase 2 logo. Therefore, if an IPv6 implementation wants to obtain an IPSec phase 2 logo, it will need to pass the IPv6 Core Protocols test in addition to the IPSec test with specific scenarios.

IPv6 Ready Logo program makes interoperability test scenarios, and conformance test tools for performing conformance tests for these corresponding certification official logos. These tests specifications are regularly updated and even though they are still considered experimental, the latest version is always available on the IPv6 Ready Logo Program portal for downloading.

The third phase was scheduled to begin in 2006 and the goals were to be the same as phase 2 but with the support of the IPSec features being an obligatory requirement. This phase was scheduled to be operational in 2006 but in 2010 is still waiting approval.

The IPv6 Ready Logo program management

Technical aspects of the IPv6 Ready Logo Program are managed by the IPv6 Program Committee (v6PC) which is composed of the JATE (Japan) [125], UNH-IOL (US) [198], IRISA/INRIA (France) [107], ETSI IPv6 Plugtests (Europe) [57], TTA (Korea) [191], BII (China) [11] and NICI v6Lab (Taiwan) [189]. The mission of the v6PC is to establish specifications for the IPv6 conformance and interoperability tests and related protocols, to develop the testing tools and to define the procedures to be provided to the IPv6 Ready Logo Program. The main contributor to the IPv6 Ready Logo Program tests specifications so far has been TAHI. However, all the test specifications are developed by the v6PC as a global team, with the combined efforts of different countries. Therefore the work is done with the resources from different teams and is technically independent from any vendor or state.

Besides the IPv6 Forum, some of the specifications are promoted by the Certification Working Group [106] of the IPv6 Promotion Council (Japan). This Certification Working Group was created in March 2003 and it has five sub-working groups which are specialized in Mobile IPv6, IPSec, SIP, DHCPv6 and IPv6 core protocols.

2.4.2. IPv6 Testing project

The IPv6 Testing Project was launched in November 2004 and is led by the Protocol and Testing Competence Centre (PTCC) of the European Telecommunications Standards Institute (ETSI) [53] and also promoted by the IPv6 Forum and the eEurope Standardization Project for the Development of Test Specifications for IPv6 created by the European Commission's eEurope Action Plan [49]. The global objective of this project is to reduce the cost of current testing and test development. This will be achieved by providing an open and free test development framework in Testing and Test Control Notation Version 3 (TTCN-3) [60], which is a new test specification language, and by the creation of conformance and interoperability TTCN-3 test packages for four key areas of IPv6: core protocol, security, mobility and transitioning. All this work will be done in close contact with the IPv6 Ready Logo Program.

The task of the PTCC is to produce conformance testing standards and products for interoperability. It is organized in Technical Committees (TC), like the TC on Methods for Testing & Specification (TC MTS). When the work is not particularly suited to the development of a TC, ETSI contracts selected experts to carry out the technical developments which are called Specialists Task Forces (STF) in order to accelerate the standardization process in response to urgent market needs. Two STFs have been created in order to develop the IPv6 test specifications.

The STF256 was in charge of the creation of a TTCN-3 framework for IPv6 test development [60] in order to provide a set of broad guidelines for the use of a common method of developing Ipv6 test specifications from November 2003 to September 2004. This toolkit is applicable to various types of IPv6 testing, including interoperability testing by the STF276.

The STF276, also known as the “IPv6 Testing Project”, is the working group on “Test specifications for IPv6 interoperability and framework for testing for interoperability” and it is divided into two sections. The first phase of the project, which began in November 2004 and finished in May 2006, had two work packages. WP1 which related to the framework of IPv6 testing for interoperability [61] and WP2 which provided the conformance and interoperability test specifications and the library of TTCN-3 test code for IPv6 Core Protocols.

The second phase of the project had four work packages related to IPv6 Security, IPv6 Mobility, IPv6 Transitioning and dissemination of results. This phase began in January 2006 and finished in July 2007.

The specifications of STF276 [61] are based on the "MTS IPv6 Testing Methodology and Framework" developed by STF256 [60] and streamline the methodologies specified by ISO [109] for conformance testing. Furthermore they apply major parts of TIPHON [62] for the interoperability methodology and are written according to the guidelines of [56] proposed by ETSI.

Many publications have proposed formal languages for IPv6 testing [195][220], but the TTCN-3 is a new standardized language for the definition of test specifications for a wide range of computer and telecommunications systems such as IPv6. TTCN-3 has been developed and standardized by ETSI [54] and has been adopted as part of the ITU-T Recommendations [121]. Syntax allows the concise description of test behaviour by unambiguously defining the meaning of a test case pass or fail. Even though protocols are the typical area of application of TTCN-3, it is not restricted to conformance testing and can be also used in areas like interoperability testing.

2.4.3. IPv6 testing tools

A testing tool, or test suite, is the software bundle that performs conformance and interoperability tests. It is usually formed by a software bundle composed of a conformance test tool (test platform), a package of IPv6 conformance tests (test scripts) and the remote files which make up for differences between test target devices. Even though a hardware test platform may be used in commercial test suites, the main test platform employed to date to perform test specifications executions has been TAHI's v6eval. However, the IPv6 ETSI's TTCN-3 substituted this testing product by 2008.

V6eval is a free tool designed for advanced IPv6 packet generation and analysis that needs to run on a FreeBSD operating system. The methodology followed to do the tests is usually based on sending customised packets and checking the response obtained. This tool is analyzed more deeply in another section.

At present, many IPv6 test suites are available (Table 3) besides the IPv6PC test suites previously discussed and many of them are publicly available and developed by open-source initiatives. There are not many commercial test suites in the market, probably due to the great development work carried out by the institutions mentioned above and by the IPv6 Ready Logo Program itself, which has provided relatively easy to use open and free tools. A notorious case was that of Navtel Communications Inc. which on acquiring NetTest Inc., with quite a complete set of test suites, decided to stop supporting IPv6 conformance testing.

The TAHI project is the pioneer in IPv6 test suites. Besides the fundamental participation in most of the v6PC specifications, it has also released some independent test packages such as NEMO (Network Mobility) or SIP (Session Initiation Protocol) for IPv6 and its own conformance test package. TAHI's test platform, v6eval, is a packet oriented test tool which is very useful for lower-layer testing, but it needs to emulate all the packets and depends on

FreeBSD. TAHI has been developing application layer protocol tests since autumn 2006. This communication oriented test tool is named Koi and is an upper layer protocol conformance test suite.

Organizations	*Test Suites*		Test Plattform	Availability
	InterOperab.	Conformance		
TAHI	✓	✓	SW (v6eval)	Free
IRISA	✓	✓	SW (v6eval)	Free
ETSI	✓	✓	SW (TTCN-3)	Free
UNH-IOL	✓	✓	SW (TestMonkey)	Free/Commercial
Agilent Technologies		✓	HW	Commercial
IXIA		✓	SW (IxANVL)	Commercial
Net-O2 Technologies		✓	SW (ATTEST)	Commercial
Spirent Communications		✓	HW	Commercial

Table 3: IPv6 test suites

The “Tests de Conformité & d'Interopérabilité des Protocoles Internet” group at IRISA is the French representative of the IPv6 Ready Logo Program certification group. Besides the v6PC specifications, it has developed some specifications of its own in a variety of IPv6 technologies. This team also collaborates in the STF276 promoted by the ETSI’s PTCC previously discussed.

The IPv6 testing at the UNH-IOL is performed by the IPv6 Consortium, which is formed by companies interested in reducing the marketing time for its members and the time needed to spread the adoption of this technology. This group offers its own test specifications focused mainly on routing technologies, besides the IPv6 Ready Logo Program specifications, and has developed a proprietary software test tool to test IPv6’s conformance testing and IPv6 Ready Logo Program functionalities called TestMonkey.

Each commercial solution has its own conformance test suites. Two of them have a specialized hardware device [1][184] and the other two have software that runs on Linux OS (IxANVL also on Windows) [124][152].

2.4.4. Interoperability testing events and laboratories

In order to evaluate the interoperability between a network device under testing and others, the IPv6 Ready Logo Program methodology requires tests with at least 4 devices from other vendors (2 hosts and 2 routers). There are two ways of obtaining the IPv6 Ready logos; taking the devices to the interoperability events or sending them to the test laboratories. The test procedures for the approval of logo usage must be exactly based on the test in both of these processes.

There are some IPv6 Ready test laboratories that can provide both conformance and interoperability testing for IPv6 devices. These laboratories are approved by the v6PC and are strategically distributed around the world. They are TTA (Korea), BII (China), NICI v6Lab (Taiwan), IRISA/INRIA (France) and UNH-IOL (US).

The IPv6 interoperability events, which tend to last about a week, take place regularly and are the place where vendors can meet and apply for any of the IPv6 Ready logos. Interoperability events are very complex and require the coordination and collaboration of several groups, most of them discussed in this chapter. Participation in interoperability events is open to all vendors and is not mandatory, as tests can also be done in the testing laboratories.

The ETSI organizes IPv6 Plugtests events which are open to all companies, organizations and work and study groups implementing the standard. The first international event was held in France in 2000, and since then Plugtest events have taken place every autumn in different locations worldwide. They are supported by IPv6 Forum, IRISA, PTCC and TAHI. The first beta Conformance Tests for IPv6 in TTCN3 testing language resulting from the IPv6 Testing Project were ran at the 6th IPv6 Plugtests in October 2005 in Sophia Antipolis (France) ETSI's home.

The annual TAHI IPv6 interoperability test event has been held in January in Chiba City (Japan) in recent years with the exception of the 7th TAHI IPv6 interoperability event which was held in association with the 6th ETSI Plugtests. These events are organized by TAHI and usually supported by IPv6 Forum, IRISA, PTCC, NICI, NEC Corp. and IXIA.

The Moonv6/UNH-IOL IPv6 interoperability test event is organized and sponsored by the UNH-IOL and takes place every autumn mainly at that university. These events always include test plans from multiple network operators, UNH-IOL, the JITC and participating equipment vendors.

2.4.5. Logo acquisition

In order to become certified with the IPv6 Ready logo, applicants must first perform and succesfully pass both the conformance test and the interoperability test and then submit an application form to v6PC with the test results. There exists a single application submission point to cover the whole world [104]. If the test results pass the global examination, applicants are issued with the correspondent logo.

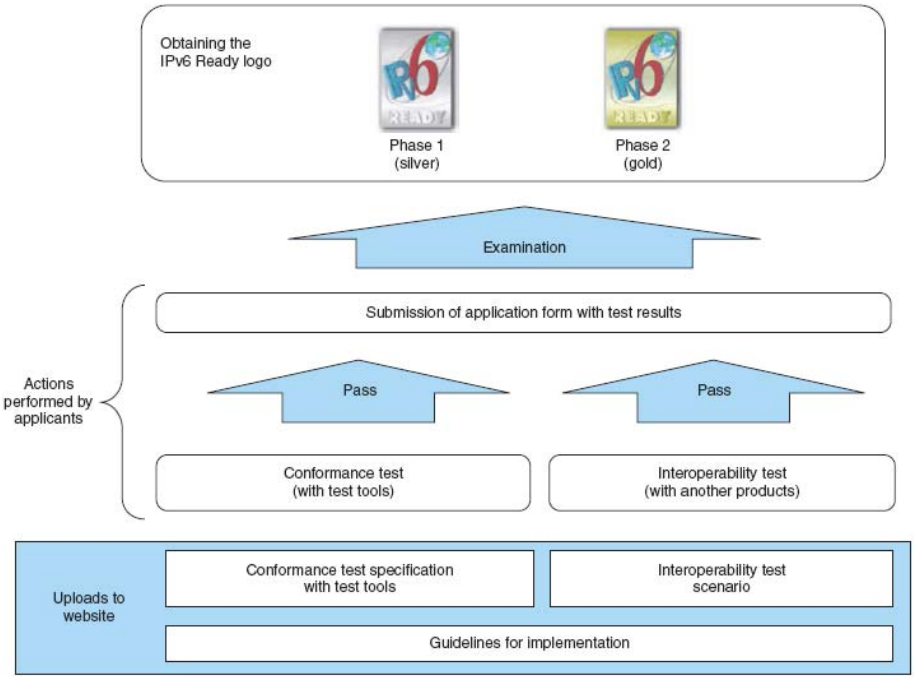

Obtaining the IPv6 Ready logo
Phase 1 (silver)
Phase 2 (gold)
Examination
Submission of application form with test results
Actions performed by applicants
Pass
Pass
Conformance test (with test tools)
Interoperability test (with another products)
Uploads to website
Conformance test specification with test tools
Interoperability test scenario
Guidelines for implementation

Figure 8: Flow chart of IPv6 Ready logo acquisition [190]

2.4.6. IPv6 Ready Logo program certified implementations

In December 2006, a study [204] was realized by the author about the certified implementations with the aim of providing a quantification of the IPv6 Ready devices and their geolocalization. This study has been updated with the technical paper [190] from the NTT Technical Review (Global Standardization Activities section) and published in December 2009 with the title “Activities of IPv6 Ready Logo program”. This latter paper provides information about the IPv6 Ready logo program from the same point of view as the study [204] in 2006. Therefore, it is relevant for this thesis to show the results from [190] in order to update and compare the provided data from 2006.

Even though many IPv6 implementations have already passed IPv6 Ready Logo phase 1 and phase 2, many other existing implementations have not been subjected to IPv6 conformance and interoperability tests, making it difficult to measure their interoperability capacity. The official list of devices that have already passed one of the IPv6 Ready Logo Program phases is publicly available on the program web page.

State of the art in 2006

Table 4 shows the geographic distribution of the phase 1 approved products according to the official list in December 2006. A total of 258 different certified products from 114 different vendors had successfully passed IPv6 Ready Logo Program phase 1. This table shows the great interest in IPv6 development in Asia. The increase of Chinese products and vendors since beginning of year 2005 has been noteworthy.

Region	Country	N. of Vendors	N. of Products	Summary by Region
Asia	Japan	32	117	205 products from 85 vendors
	Taiwan	22	37	
	Korea	12	29	
	China	14	17	
	India	4	4	
	Phillippines	1	1	
North America	USA	20	35	37 products from 22 vendors
	Canada	2	2	
Europe	Germany	1	2	7 products from 5 vendors
	Sweden	1	2	
	Austria	1	1	
	Denmark	1	1	
	France	1	1	
Oceania	New Zealand	1	8	8 products from 1 vendor
Middle easr	Israel	1	1	1 product from 1 vendor
		114	*258*	

Table 4: Phase 1 IPv6 Ready logos issued in 2006

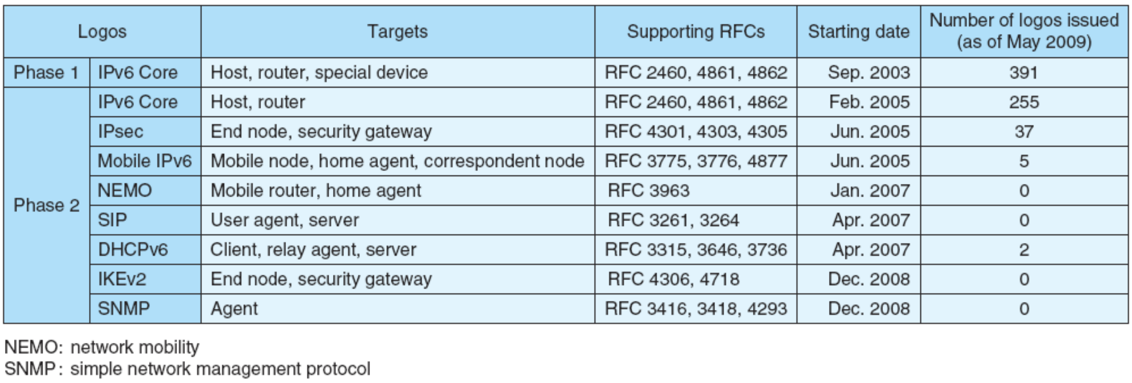
In December 2006, some aspects of IPv6 Ready Logo Program phase 2 were still under development, but the tests for MIPv6, IPSec and IPv6 core specifications were available. Table 5 shows how 55 products from 42 vendors had already obtained the IPv6 core specifications logo, with 12 of these products having also obtained the IPSec logo and 4 of them the MIPv6 logo. Table 5 also shows the geographic distribution for phase 2 approved products in December 2006.

Region	Country	N. of Vendors	N. of Products	Summary by Region
Asia	Japan	10	16	37 products from 26 vendors
	Taiwan	5	9	
	China	6	7	
	Korea	4	4	
	India	1	1	
North America	USA	13	15	16 products from 14 vendors
	Canada	1	1	
Europe	Denmark	1	1	2 products from 2 vendors
	France	1	1	
		42	*55*	

Table 5: Phase 2 IPv6 Ready logos issued in 2006

State of the art in 2009

The NTT Technical Review paper [190] shows that there has been a great evolution of the certified devices from December 2006 (Table 4 and 5) until May 2009 (Figure 9), especially in the number of issued phase 2 logos which have been increased in more than 200.



Logos		Targets	Supporting RFCs	Starting date	Number of logos issued (as of May 2009)
Phase 1	IPv6 Core	Host, router, special device	RFC 2460, 4861, 4862	Sep. 2003	391
Phase 2	IPv6 Core	Host, router	RFC 2460, 4861, 4862	Feb. 2005	255
	IPsec	End node, security gateway	RFC 4301, 4303, 4305	Jun. 2005	37
	Mobile IPv6	Mobile node, home agent, correspondent node	RFC 3775, 3776, 4877	Jun. 2005	5
	NEMO	Mobile router, home agent	RFC 3963	Jan. 2007	0
	SIP	User agent, server	RFC 3261, 3264	Apr. 2007	0
	DHCPv6	Client, relay agent, server	RFC 3315, 3646, 3736	Apr. 2007	2
	IKEv2	End node, security gateway	RFC 4306, 4718	Dec. 2008	0
	SNMP	Agent	RFC 3416, 3418, 4293	Dec. 2008	0

NEMO: network mobility
SNMP: simple network management protocol

Figure 9: Number of IPv6 Ready logos issued in 2009

In Figure 10 it can be seen the number of acquired phase 2 golden logos by country until May 2009. If this data is compared with the data from Table 5, it is shown how USA and Japan have largely increased their phase 2 certificated devices (from 15 and 16 respectively in 2006 to 104 and 69 respectively in 2009). It is surprising, though, the slow advance in this field of the Chinese certificated products as it has just increased from 7 to 28.

In August 2008, the National Institute of Standards and Technologies (NIST) released “A Profile for IPv6 in the U.S. Government ver. 1.0” [157] in order to develop the technical infrastructure (standards and testing) necessary to support wide scale adoption of IPv6 in the US Government (USG). This profile is based on the IPv6 Ready Logo Program specifications. This is why the number of phase 2 logos acquired in the USA has grown so remarkably in the last two years.

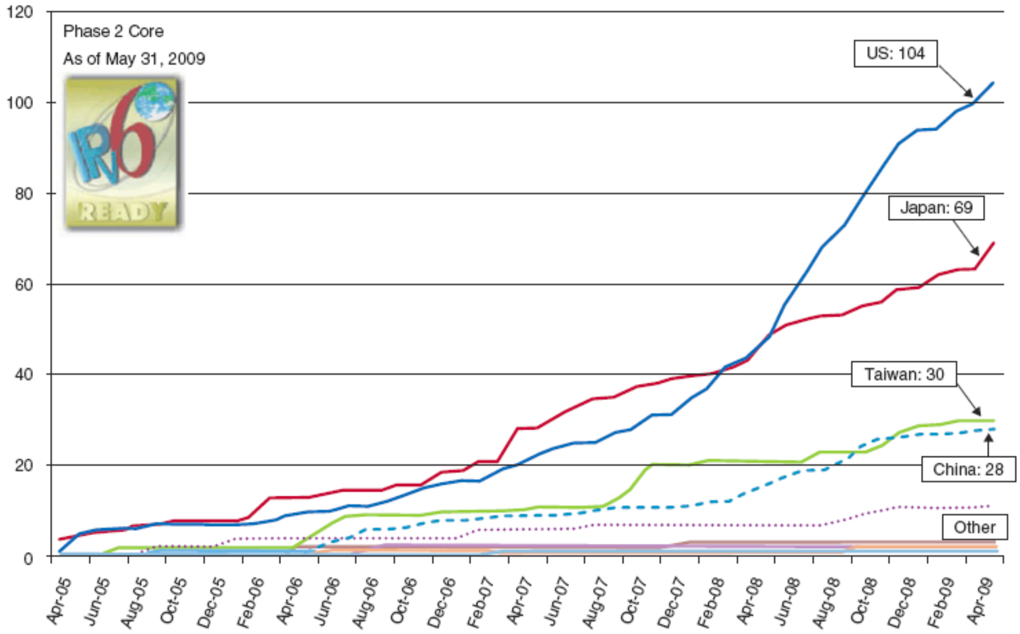


Figure 10: Phase 2 IPv6 Ready logos issued in 2009 by country [190]

2.5. Conformance tests and results

In year 2004 it was decided to make IPv6 conformance tests over different end node operative systems allegedly supporting IPv6 with the final aim of implementing an IPv6 testbed. The idea of the testbed was to use computers as routers with some performance guarantees and to develop researching over the testbed.

When the testing took place, the IPv6 Ready Logo program only certified NUTs with the phase 1 (silver) logo. It must be noticed, though, that phase 1 provided a good approach, even though it does not guarantee fully conformance to IPv6 core specifications. At that time the IPv6 Ready Logo program did not use to publish explicitly the phase 1 certified devices as it does currently. The only way to get that information was directly from the vendors (rarely published), from the TAHI project site (several months outdated) or from the KAME project site, which offered very synthetic updated information but uniquely from its action field.

The state of the arts showed very few interesting similar studies [141][228], and these were very limited and did not offered any comparative studies useful for our project.

Therefore the study presented in this thesis provided in 2005 [178] an innovating comparative study of the conformance level of the different IPv6 supporting operative systems and had been certified by the IPv6 Ready Logo program phase 1.

2.5.1. Testing enviroment

It was decided to use the TAHI project's test suite (full CT package) instead of the suite provided by the IPv6 Ready Logo program phase 1 or phase 2 (beta) because the the former was much more complete than the latter and the objective was to evaluate functionalities not included in any of the IPv6 Ready Logo program phases, like the PMTUD specification.

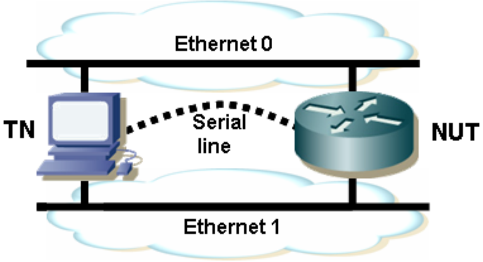


Figure 11: TAHI Conformance Test scenario

The TAHI Conformance Test Suite is a bundle of software composed by a Conformance Test tool called v6eval and a package of IPv6 conformance tests called CT (in script form). V6eval is a very powerful tool designed for advanced IPv6 packet generation and analysis. V6eval presumes a test environment with a Tester Node (TN) directly connected to the Node Under Test (NUT) via one or multiple Ethernet connections, depending on the type of NUT. This environment is shown in Figure 11. During the testing process the TN sends packets to the NUT and checks for responses. Sometimes, however, packets must be originated by the NUT. In this case, input/output command line shall be required on the NUT. In order to fully automate the testing process, a RS-232 serial line can be attached between the TN and the NUT, and then commands can be sent from the TN interactively. V6eval supports remote control scripts for more than 30 different IPv6 implementations in order to automate the testing process. The last release of v6eval analyzed by the author was version 2.3.2, released on January 2005. CT is the script package that contains the main branch of TAHI tests. The last release of CT analyzed by the author was version 2.1.2 and was updated for the last time on February 2005. This package was the most complete set of tests released by TAHI until then. The following table shows the tests supported by the CT package:

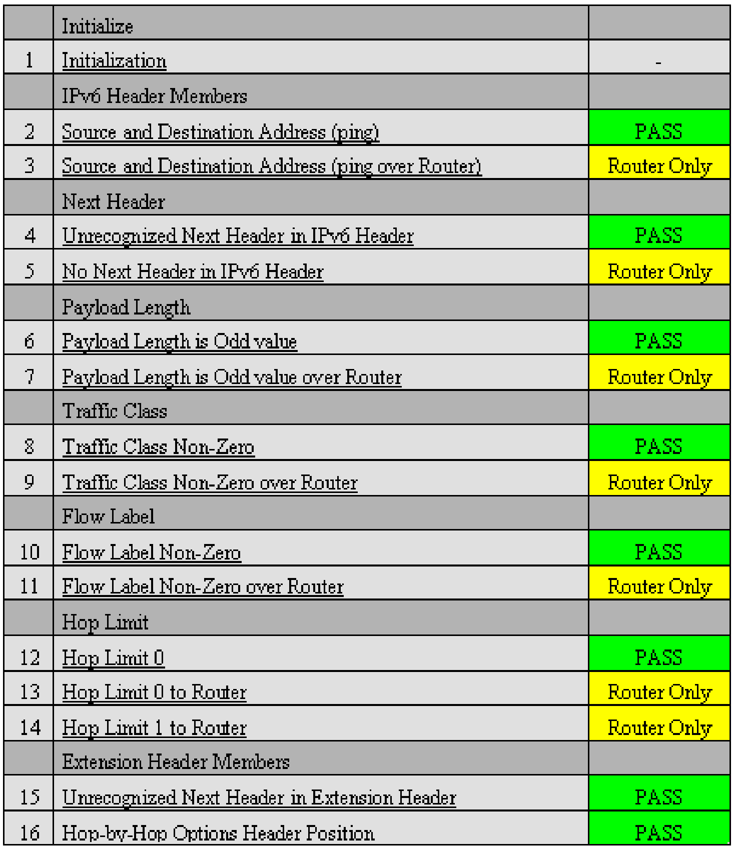
- IPv6 core Specification
- ICMP for IPv6
- Neighbor Discovery for IPv6
- IPv6 Stateless Address Autoconfiguration
- Path MTU Discovery for IPv6
- IPSec
- Mobility Support in IPv6
- Transition Mechanisms for IPv6
- IPv6 Prefix Options for DHCPv6
- DNS Discovery
- Default Address Selection for IPv6
- Stateless IP/ICMP Translation Algorithm (SIIT)
- Network Address Translation – Protocol Translation (NAT-PT)
- IPv6 Router Selection
- IPv6 Protocol Stack Robustness

The Robustness test is the only one not based on any RFC or Draft. It is a custom test designed for network stress testing [188]. Since July 2003, TAHI also releases some independent test packages such as MIPv6, SIP or the IPv6 Ready Logo test series.

The testing evaluated the developing status of different host implementations. It was decided to evaluate only host devices because the available TAHI project's CT package was only prepared for host testing. The testing for router devices was published afterwards. The v6eval tests can result in one of the following cases:

- **PASS:** No problems were found while running the test.
- **FAIL:** The test was not performed correctly.
- **WARN:** In general it may be defined as an expected error.

TAHI has always evolved its packages and therefore the format of the results. Initially the results were saved in a text file, in 2005 evolved and since then are presented in HTML format. The results of this newer form of presenting results is shown in Figure 12.



	Initialize	
1	Initialization	-
	IPv6 Header Members	
2	Source and Destination Address (ping)	PASS
3	Source and Destination Address (ping over Router)	Router Only
	Next Header	
4	Unrecognized Next Header in IPv6 Header	PASS
5	No Next Header in IPv6 Header	Router Only
	Payload Length	
6	Payload Length is Odd value	PASS
7	Payload Length is Odd value over Router	Router Only
	Traffic Class	
8	Traffic Class Non-Zero	PASS
9	Traffic Class Non-Zero over Router	Router Only
	Flow Label	
10	Flow Label Non-Zero	PASS
11	Flow Label Non-Zero over Router	Router Only
	Hop Limit	
12	Hop Limit 0	PASS
13	Hop Limit 0 to Router	Router Only
14	Hop Limit 1 to Router	Router Only
	Extension Header Members	
15	Unrecognized Next Header in Extension Header	PASS
16	Hop-by-Hop Options Header Position	PASS

Figure 12: Report of a TAHI project conformance test

The NUT devices chosen for the test were:

- Microsoft Windows 2003 Server
- Linux (kernel 2.6.6)
- Linux + USAGI *snap* linux26-s20041011
- FreeBSD 4.9
- FreeBSD + KAME *snap*

It must be noticed the differences between the two Linux NUTs and the two UNIX NUTs. The 2.6.6 Linux kernel version already incorporated the stable USAGI patch for IPv6. Actually, Linux has been incorporating it since the first publication of the 2.6 version. The other Linux

NUT used for testing had incorporated the latest regularly published stable version of the USAGI patch, named snap. The same way as with Linux, FreeBSD also incorporates an stable KAME patch since version 4.0. For the testing, and similarly to what was done with the two Linux versions, it was intended to compare the official version of the O.S. with the latest stable snap released by KAME. It also was intended to compare with the latest Microsoft Windows version released at that time which was the Microsoft Windows 2003 Server. This O.S. was the first Microsoft O.S. to incorporate natively IPv6.

The TN node has installed a FreeBSD 4.9 O.S. in order to run the v6eval software. All the NUTs were tested under the same identical conditions and the tested O.S. were installed over identical Pentium IV 2.4 GHz with 512Mb of RAM machines.

2.5.2. Results

The test results are shown in Figures 13, 14 and 15, where it can be observed a comparative study of the five O.S. analysed in three significant tests:

- IPv6 specification (RFC-2460) [37]
- ICMPv6 specification (RFC-2463) [31]
- PMTUD specification (RFC-1981) [137]

The first two tests were partially included in phase 1 of IPv6 Ready Logo program at that time, and now they are totally included. It must be noticed that, currently, phase 1 supports RFC-4443 instead of RFC-2463 which is obsolete.

The results for the five implementations were very good in the first two tests. Actually, the Linux performance was outstanding within both versions of the O.S., the official and the version with the snap. A pretty good test was also performed by the UNIX O.S. with the latest KAME patch but conformance test showed some few lacks in the implementation of the IPv6 specification. Official UNIX O.S. and Windows 2003 had the poorest results, especially the latter.

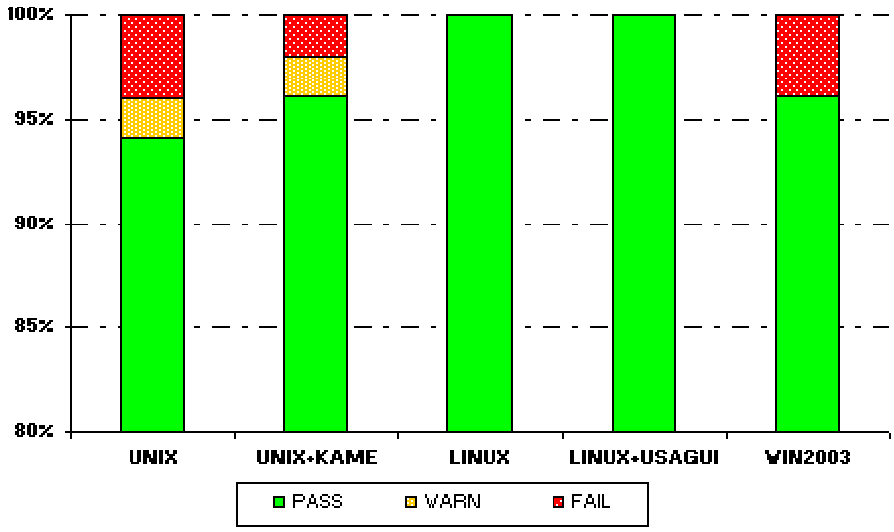


Figure 13: Test results of the IPv6 specification conformance test

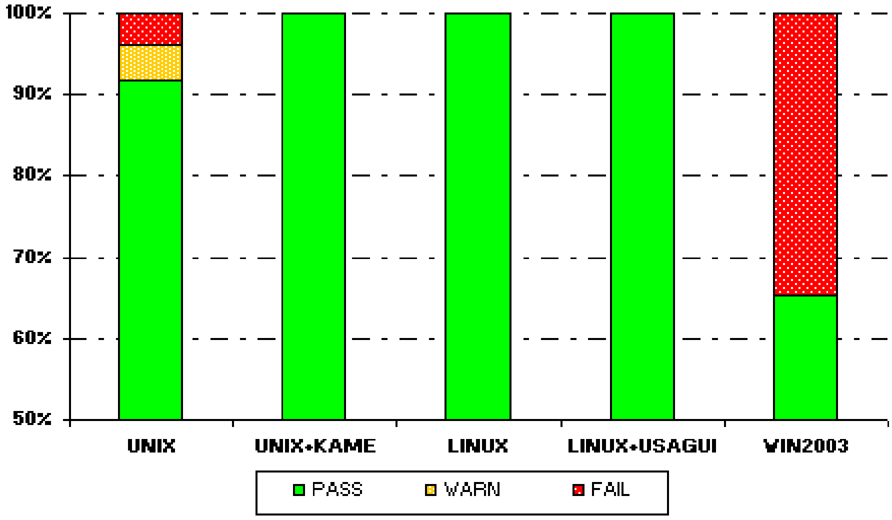


Figure 14: Test results of the ICMPv6 specification conformance test

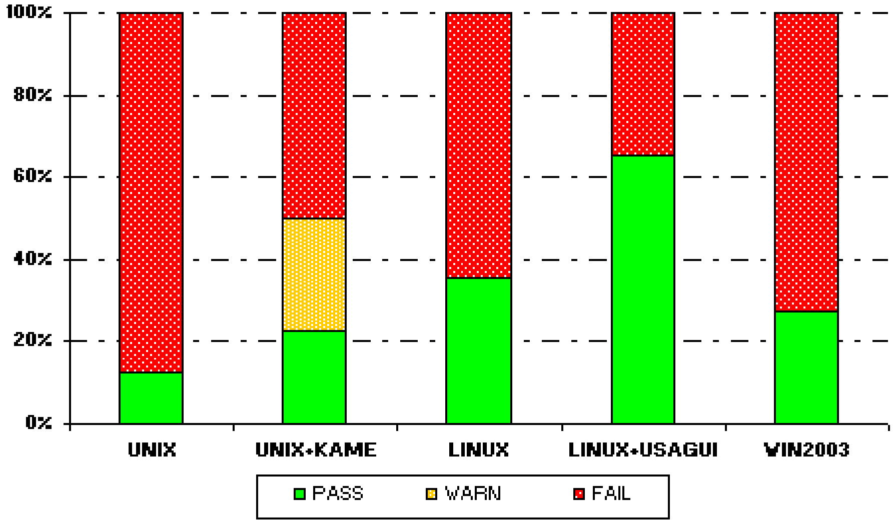


Figure 15: Test results of the PMTUD specification conformance test

The PMTUD specification was not intended to be proved on phase 1. Even though the results showed that it is one of the fewest features developed by the tested end node O.S.. It was decided to test the PMTUD specification because it was under discussion its compulsoriness for phase 2. Currently, it is included as part of the phase 2 IPv6 core specifications. The results show how the USAGI snap was the most standardized, even though not completely.

2.5.3. IPv6 testbed

The practical application of these test results were the implementation of an IPv6 testbed, stable and open source, as it can be seen in Figures 16 and 17. The main routers (Homer, Marge and Maggie) were computers with Linux with Debian distributions and USAGI patches. This tesbed had connection to the 6bone through the Cesca and Rediris networks. Later on this testbed was updated to support QoS as it is explained in the next section.

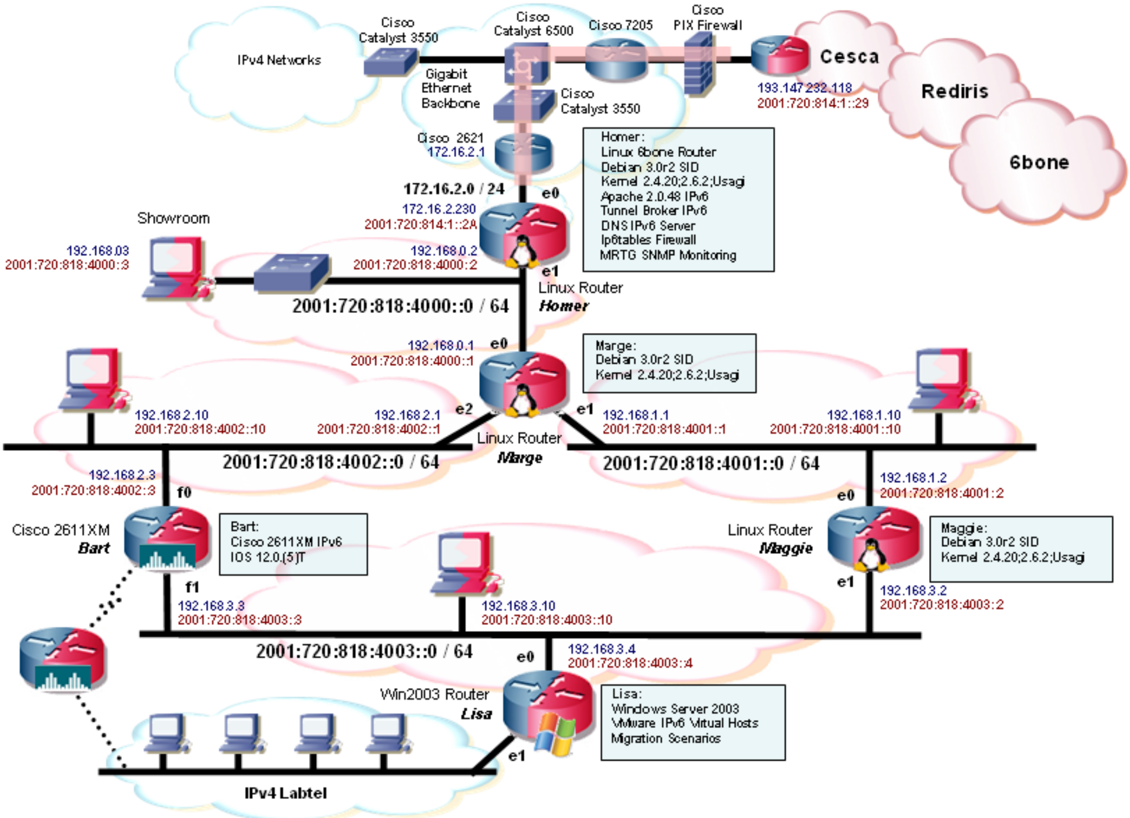
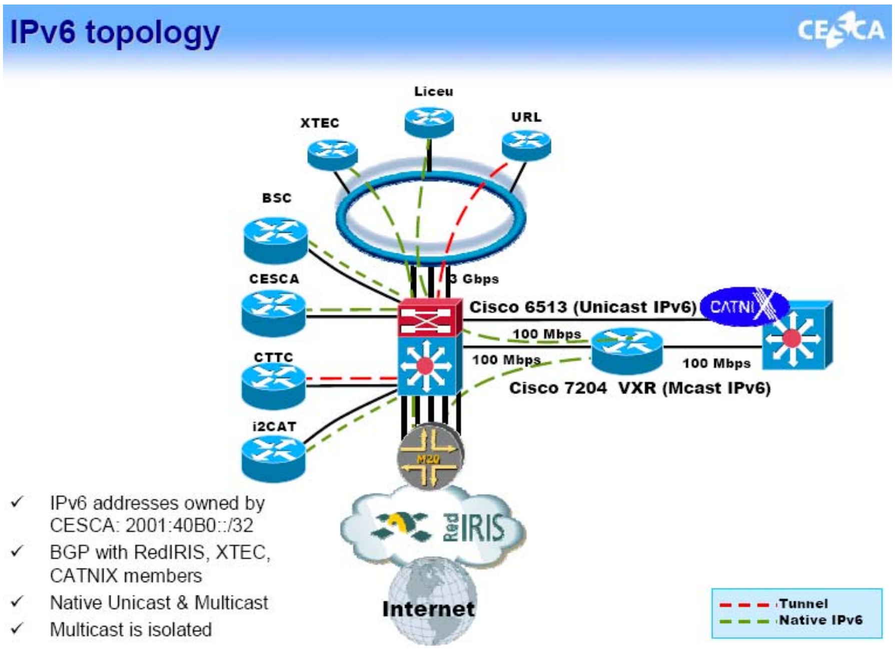


Figure 16: First version of the IPv6 testbed in the Networking Laboratory – URL

Figure 17: Image of the IPv6 testbed in the Networking Laboratory

In Figure 18 it is illustrated a detail of the Cesca connetion to the Internet through the RedIris network, including the Universitat Ramon Llull (URL) IPv6 tunnel connection indicated in red in Figure 18 and in pink in Figure 16. This connection is an IPv6 tunnel over the regular IPv4 network.



- ✓ IPv6 addresses owned by CESCA: 2001:40B0::/32
- ✓ BGP with RedIRIS, XTEC, CATNIX members
- ✓ Native Unicast & Multicast
- ✓ Multicast is isolated

Figure 18: CESCA's IPv6 topology

This IPv6 testbed has been used as a platform for several of the La Salle/URL alumni projects, including a tunnel broker application for providing IPv6 Internet connection to any IPv4 computer from La Salle/URL, as shown in the IPv6 testbed gateway in Figure 16.

On July 28th, 2009, the IPv6 testbed was definitively disconnected as part of the moving to a newer laboratory. There is no further intention of rebuild it by the new lab managers. At present, there is no IPv6 connection to the Internet from La Salle/URL.

2.6. Chapter summary

The IPv6 protocol should have been implemented in most of the world IP devices and networks but it has not been. Transition is still in on its way even though IPv4 addresses will have run out by 2011. This chapter has reviewed the main stakeholders in promoting and developing the IPv6 and the most important conformance and interoperability initiatives which provide quality assessment for existing IPv6 implementations, through a detailed study of the IPv6 Ready Logo Program and the ETSI IPv6 Testing Project. These can be carried out autonomously with the existing IPv6 testing tools or with the IPv6 interoperability testing options at the interoperability test events that periodically take place around the world. Both the IPv6 Ready Logo Program and the ETSI's IPv6 Testing Project are already operational, and are finalizing the final phases of their conformance testing projects.

All these tools allow the practical deployment of this technology with guarantees. The IPv6 conformance test carried out proved that the three main end node O.S. had excellent results back in 2005 and demonstrated the great work carried out by the Japanese WIDE project. The tests analysed three basic IPv6 specifications for each O.S. These contributions to the thesis have been presented in [178] and [204]. Moreover, these first two scientific papers approved by the international scientific community validated the quality of the research methodology of the author.

Two main conclusions can be extracted from the overall analysis of this chapter. The first conclusion is that after ten years of work all the necessary building blocks for the successful and massive worldwide deployment of IPv6 are in place and that the interest in IPv6 deployment has increased on a global scale. Thus, it can be concluded that IPv6 technologies have entered the final phase of real practical usage and massive deployment and that now is the key moment for the definitive boost and massive deployment of IPv6.

The second conclusion is that, GNU/Linux based systems with the USAGI snap performed outstanding IPv6 conformance tests and was the most standardized of the tested end node O.S. Therefore this end node O.S. is ready to be used as a base for a stable IPv6 testbed with performance guarantees from which new research lines may be opened.

3. AUTOMATED QOS MANAGEMENT IN BROADBAND NETWORKS

3.1. Introduction

The usage of IP technologies is a common feature in core and access areas of the NGNs which need QoS requirements to support their services and where the need to support IPv6 is one of the basic premises.

The importance of the IPv6 protocol in current communication networks has been already stated in this thesis. IPv6 is currently supported by the main software and hardware applications, which have been already certified, and it is present in the main worldwide networks. Moreover, as a result of the work carried out by the USAGI project, which has merged its work into the official Linux kernel, the IPv6 stack in Linux OS is now fully compliant with advanced IPv6 conformance and interoperability tests. Thus, the implementation of real IPv6 testbeds with Linux OS with performance guarantees has been possible since June 2006.

The QoS management for the new IP technologies in NGNs has undergone important advances with the introduction of the Policy-based network management (PBNM) for resource allocation. NGNs require automatic provisioning of QoS and PBNM simplifies the definition and deployment of network policies through centralized management frameworks. Therefore, the centralized management of networks with QoS and/or Traffic Engineering has become a fundamental issue in recent research. Some proposals are the Bandwidth Brokers (BB) [155] for DiffServ networks or the Path Computation Elements (PCE) [97] for Multi Protocol Label

Switching networks with Traffic Engineering support (MPLS-TE).

The bandwidth brokers are devices which are capable of automatically managing the QoS using a centralized architecture within a DiffServ domain. These devices have been proposed for the control and management of QoS provisioning to reduce the complexity of the control plane. They are basically resource controllers which manage the limited amount of resources specified by the client contracts or the Service Level Agreements (SLAs) of a DiffServ domain and make the service allocation decisions from the resources to be applied to the nodes of that domain. The bandwidth brokers allow the administrator to configure network policies with a high level language and a friendly interface and save these policies in a structured way at the different nodes involved. They are then configured within these nodes using specialized policy-based transmission protocols, making them transparent to the administrator. These architectures are able to manage DiffServ networks using an intra-domain communication protocol for policy delivery and are able to interact with other peer bandwidth brokers in other DiffServ domains by using an inter-domain communication protocol.

Other examples are the PCEs, which are defined by the Path Computation Element Working Group at the IETF (PCE WG) [97] for managing the MPLS-TE and the MPLS DS-TE (DiffServ-Aware MPLS-TE) [63] core networks. This working group states that there is a need for a centralized architecture to manage the computation of paths of MPLS and Traffic Engineering LSP paths, as well as the QoS of the different domains, and to perform the delivery of the network policies to the nodes. The PCE WG proposes COPS-PR (Common Open Policy Service for Policy Provisioning) [23] for policy delivery.

This chapter implements the proposal of the CSI-RHET project to implement a bandwidth broker architecture in order to manage DiffServ QoS. One of the predicted results of this application was to be further enhanced in order to manage LSPs in MPLS DS-TE networks.

3.1.1. Objectives

The global objective of this second line of research was to implement a bandwidth broker architecture for the centralized management of the QoS of dual IPv4 and IPv6 networks as stated in the CSI-RHET project. In order to develop such architecture, two aspects must be taken into account: policy creation and policy application. On the one hand the system administrator must be able to introduce the QoS policies in an organized, reliable, coherent and secure way, and it must also be able to store these policies in a data base. On the other hand, a communication protocol is needed to deliver policies to the domain nodes (mainly routers) in a secure and optimized way, and these policies have to be applied in the nodes efficiently. A communication protocol for interaction between peer bandwidth brokers is also required.

The first objective was to design and implement a bandwidth broker device whose design had to have a policy framework, a friendly and secure interface for a remote administrator and a standardized policy format. The election of the communication protocols for the policy delivery was also a key aspect to take into account.

The second objective was to empirically demonstrate the practical uses of the proposed architecture. Therefore a multi domain testbed had to be designed and implemented in order to evaluate the correct performance of the whole policy-based system, including inter-domain communications. The testbed had to be implemented with computers acting as routers with GNU/Linux O.S., with IPv6 support and with QoS mechanisms. The success of this second phase of the thesis will be measured by the correct functioning of the bandwidth broker architecture, built according to the aforementioned requirements.

3.2. QoS in data networks

One of the basic requirements for the converged IP networks is the ability to provide real-time services, typically voice, with minimum services guarantees making user QoE similar to that provided by the PSTN. Regular IP networks, commonly known as Internet 1.0, cannot provide any QoS as they are best effort services by nature. Therefore it must be enhanced in order to achieve the QoS needed in the NGN. The data packets must behave in a consistent and predictable way in terms of bandwidth (throughput), latency, jitter and packet loss. Therefore the QoS refers to the capability of a network to provide better service to selected network traffic over various technologies.

Certain critical applications for companies will have to be prioritized. These include Telephony over IP (ToIP), videoconference or any traffic which is vital for company survival. The QoS provides added value through a contract with the service provider or Service Level Agreement (SLA). Furthermore, the number of data networks customers has greatly increased in recent years and this has led to network segmentation. Demand is no longer homogenous and we now have different clients with different needs. From the business point of view, the possibility of offering different levels of quality to meet different customer profiles will be beneficial for service operators, as they will be able to fix prices according to QoS of each service class provided.

There are various technologies available which provide QoS in IP networks. The integrated services model (IntServ) is focused on reserving network resources based on flows. The differentiated services (DiffServ) model works by prioritizing aggregated and categorized flows. Broadly speaking, it could be said that using DiffServ model in networks offers more scalability than those using IntServ, as the former do not guarantee end-to-end resources. In the DiffServ model, QoS cannot be guaranteed and the highest aspiration is to achieve the desired QoS with maximum probability.

Alternative technologies for providing QoS do exist but they work in different layers of the ISO layered model and, therefore, they are out of the scope of this chapter. Some layer 2 technologies for managing QoS are the 802.1p and 802.1q specifications, the EXP field for the MPLS technology or the 802.11e for WLAN technology. This latter will be discussed at a later point when QoS management in BPL networks is reviewed.

3.2.1. DiffServ

The differentiated services mechanism, DiffServ, is the most implemented QoS mechanism in current routers. The main advantage of this technology is that implies very few changes in the existing IP networks and it is easy and efficient.

DiffServ networks were introduced to solve the implementation and deployment difficulties of IntServ networks. The Differentiated Services Working Group (DiffServ WG) [95] defined an architecture based on pushing complexity to the edges of the network and keeping classification and packet handling functions in the core network as simple as possible [12].

Moreover, DiffServ offers more scalability of QoS provisioning than the IntServ networks because "the amount of state information is proportional to the number of classes rather than the number of flows" [217]. Therefore, this mechanism proposes an architecture where the treatment is done on aggregated flows and it treats exactly the same way to all the flows marked with the same QoS.

The flows are classified by the MultiField (MF) classifier, and then metered, policed, marked and shaped at the edge nodes of a DiffServ domain. The core nodes handle packets according to Per Hop Behaviors (PHBs) [12] which are selected on the basis of the Behavior Aggregate (BA) classifier which selects packets based exclusively on the DS field contents. The DiffServ codepoint (DSCP) in the DiffServ Field maps the class of service in every IP packet header, IPv4 and IPv6 (Figure 19).

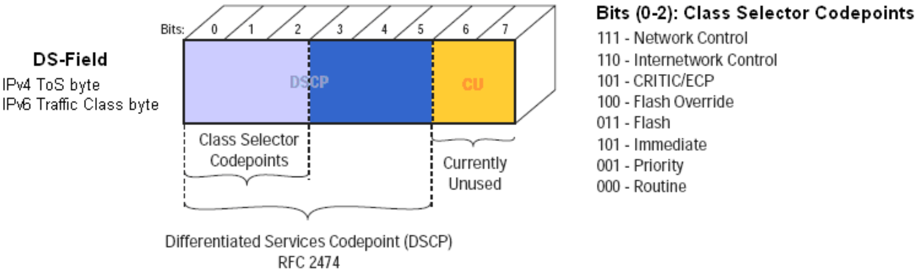


Figure 19: DiffServ codepoint (DSCP)

Therefore, the PHB is the forwarding treatment applied to a collection of packets with the same DSCP which cross a link in a particular direction at a DiffServ-compliant node. When a packet ingresses into a DiffServ domain from another DiffServ domain, its DSCP value may be re-marked according to the SLA between the two domains. There exist three kinds of priorization (Table 6): EF (Expedited Forwarding) [35] for the maximum priority real-time traffic, (e.g. voice), where it provides minimum delay and jitter, AF (Assured Forwarding) [80] which provides twelve priorities and BE (Best Effort) for the non priority traffic.

PHB	DSCP (bin)	DSCP (dec)	Priority
BE	000000	0	Global Minimum
AF11	001010	10	AF1 low
AF12	001100	12	AF1 medium
AF13	001110	14	AF1 high
AF21	010010	18	AF2 low
AF22	010100	20	AF2 medium
AF23	010110	22	AF2 high
AF31	011010	23	AF3 low
AF32	011100	28	AF3 medium
AF33	011110	30	AF3 high
AF41	100010	34	AF4 low
AF42	100100	36	AF4 medium
AF43	100110	38	AF4 high
EF	101110	46	Global Maximum

Table 6: DiffServ priorizations (PHB): EF, AF and BE

Therefore, the packets enter into a DiffServ domain with IP networks with a given mark indicating its QoS priority assigned in the domain's entrance. All the routers in the domain must provide the DiffServ mechanisms to discriminate packets according to the marked QoS priority. These mechanisms include classifiers, meters, markers, policers, droppers, queuing and shapers (Figure 20).

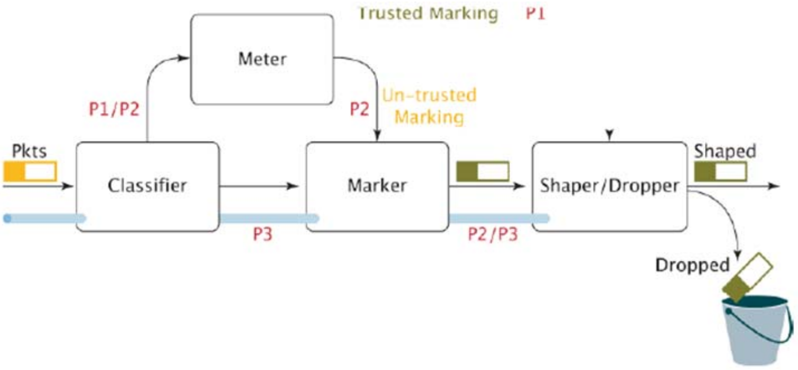


Figure 20: DiffServ mechanisms

3.2.2. Service Level Agreement

QoS negotiation gives users the ability to specify the desired level of service, along with the permissible pricing of the service. The negotiated levels are known as the Service Level Agreement (SLA). SLA management in IP networks addresses issues such as SLS parameter definition, the utility model for SLA management, and information models for policy representation.

In DiffServ, the SLA is a service contract between a customer and a service provider that specifies the forwarding service a customer should receive [12]. A customer may be a user organization (source domain) or another DS domain (upstream domain). Both must have a service contract, or SLA, with its ISP to receive differentiated services where the service classes supported and the amount of traffic allowed in each class will be specified through the Service Level Specifications (SLSs) [75]. For example, an SLS within DiffServ defines the parameters for compliant network behaviour including specific DSCPs, PHBs, profile characteristics, and treatment of the traffic for these DSCPs. The structure of an SLA is illustrated in Figure 21.

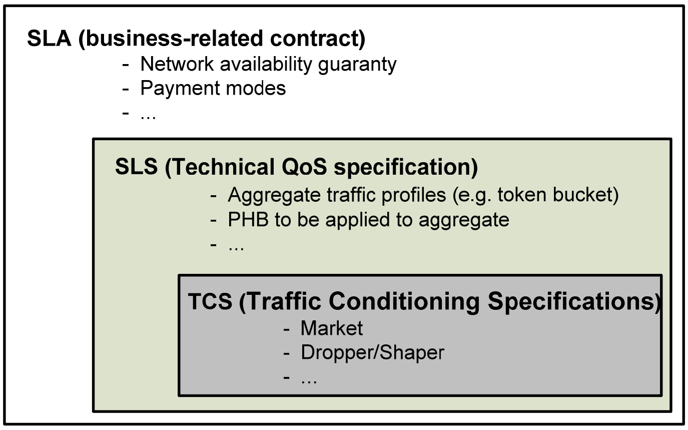

SLA (business-related contract)
- Network availability guaranty
- Payment modes
- ...
SLS (Technical QoS specification)
- Aggregate traffic profiles (e.g. token bucket)
- PHB to be applied to aggregate
- ...
TCS (Traffic Conditioning Specifications)
- Market
- Dropper/Shaper
- ...

Figure 21: SLA structure

The SLA specifies both business information and technical information.

- Business information could be price, penalty rules, the administrative information of costomer and provider.
- Technical information specifying the QoS of IP traffic is called the SLS (Service Level Specification) which defines the traffic profiles of the aggregate and the PHB to be applied to the aggregate.

A Service Level Specification (SLS) is defined as a set of parameters and their values which together define the service offered to a traffic stream by a DS domain. The SLS are composed of Traffic Conditioning Specifications (TCS) which sspecify the set of classification rules and the traffic profiles. A TCS is an integral element of an SLS.

Usually, SLAs and the included SLSs are statically negotiated for a long time period. Frequent modification is almost impossible because network configuration is manually carried out. When modification is needed, the customer may require the assistance of the network provider. The provider, in turn, may need to reengineer part of a system to ensure the mandated service levels.

3.3. QoS policy-based network management

Policy-based network management (PBNM) is an approach to configure a great number of network devices in an administrative domain to implement a set of QoS services based on IP. Basically, network operators negotiate service level agreements (SLAs) that describe the sets of QoS services they have mutually contracted to provide. Individual operators then transform the QoS requirements specified in the SLAs into sets of policy rules that are applied to their network domains to implement the contracted IP QoS services. These policy rules describe the amount of network resources required to realize the QoS services without going into the details of how to configure the network devices. The policy rules are then translated into network device configuration actions by automated entities in the domain that intimately know the network topology and the deployed network devices. Hence, the PBNM provides a high abstraction view of a network to its operator, and this helps the operator in the deployment of new QoS services as it does not need to consider details such as the size or complexity of its network. The use of automated policy translation entities further facilitates dynamic control of network resources.

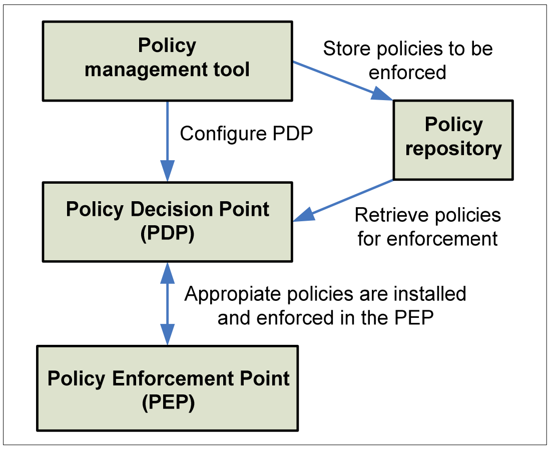

Policy management tool
Store policies to be enforced
Policy repository
Configure PDP
Policy Decision Point (PDP)
Retrieve policies for enforcement
Appropiate policies are installed and enforced in the PEP
Policy Enforcement Point (PEP)

Figure 22: PBNM architecture derived from the Policy WG framework

The need for tools to managing networks in a centralized form led the IETF to the creation of the Policy Framework Working Group (Policy WG) [98], which has provided a framework to enable the centralized control of a domain identifying the network behavior with a high level language independent of the devices and protocol that form it (Figure 22). The architecture proposed by the Policy WG has four main components: the policy management tool, the policy repository, the Policy Decision Point (PDP) and the Policy Enforcement Point (PEP). PEPs often reside in policy-aware network nodes that carry out actions stipulated by policy rules. The actions taken are based on the decisions of a PDP, which retrieves the policy rules from a repository. The PDP is the final authority the PEP needs to refer to for actions to be taken.

The Policy WG stated that an intra-domain protocol is needed for the policy decisions transmission from the PDP to the PEP within a domain, and proposes the use of Common Open Policy Service (COPS) [42] or one of its extensions.

3.3.1. Intra-domain comunication protocol

The Resource Allocation Protocol Working Group (RAP WG) [99] later became in charge of the standardization of this protocol of intra-domain communication between the policy servers (PDPs) and the network devices (PEPs) developing COPS. This protocol is a stateful query/response protocol, uses TCP (port 3288) and supports two common models for policy control: Outsourcing and Configuration.

The Outsourcing mode is the COPS protocol variation for the QoS IntServ mode. In this mode the PEP delegates responsibility to an external policy server (PDP) to make instantaneous policy decisions on its behalf. In this case the COPS-RSVP protocol is basically used to encapsulate the RSVP protocol requests and send it to the PDP [81].

In the Configuration mode of the COPS protocol, the PDP may proactively provision the PEP. This mode is defined by the Common Open Policy Service for Policy Provisioning (COPS-PR) protocol [23], which is the COPS evolution to cover the DiffServ model needs. Three separate procedures that must be performed in the COPS-PR protocol are COPS-PR initialization, PEP configuration request, and policy-changes provisioning (Figure 23). The COPS-PR initialization procedure must be completed successfully before any policy information transaction can be performed. A simple policy information transaction such as a PEP configuration request involves the exchange of a sequence of COPS-PR Request, Decision, and Report State messages.

In the COPS-PR provisioning model the PDP and the PEP each have a virtual container called PIB (Policy Information Base) where the policies are stored. There exist a different PIB for every PEP with different policies, and therefore, for each domain there will be a unique PIB for all the core routers and one PIB for each PEP router (as they need special admission policies). The PDP has a copy of all the PIBs and each PEP has a copy of its own. The PDP sends the appropriate policies to the PEP once this has been initiated and whenever there are updates. It must be noticed that it is not sent the whole PIB, only the content of the fields. Initially it must be sent all the contents but afterwards only the updated fields. This way the PDP keeps the two PIBS synchronized. These PIBs have a tree structure formed by PRovisioning Classes (PRCs) which contain PRovisioning Instances (PRIs) [138].

The first advantage of using PIB is that the protocol becomes independent of the policy information carried. This characteristic avoids revisiting the protocol every time new information needs to be carried by the protocol. The second advantage is that the classes defined in a PIB are reusable or extendable in other PIBs.

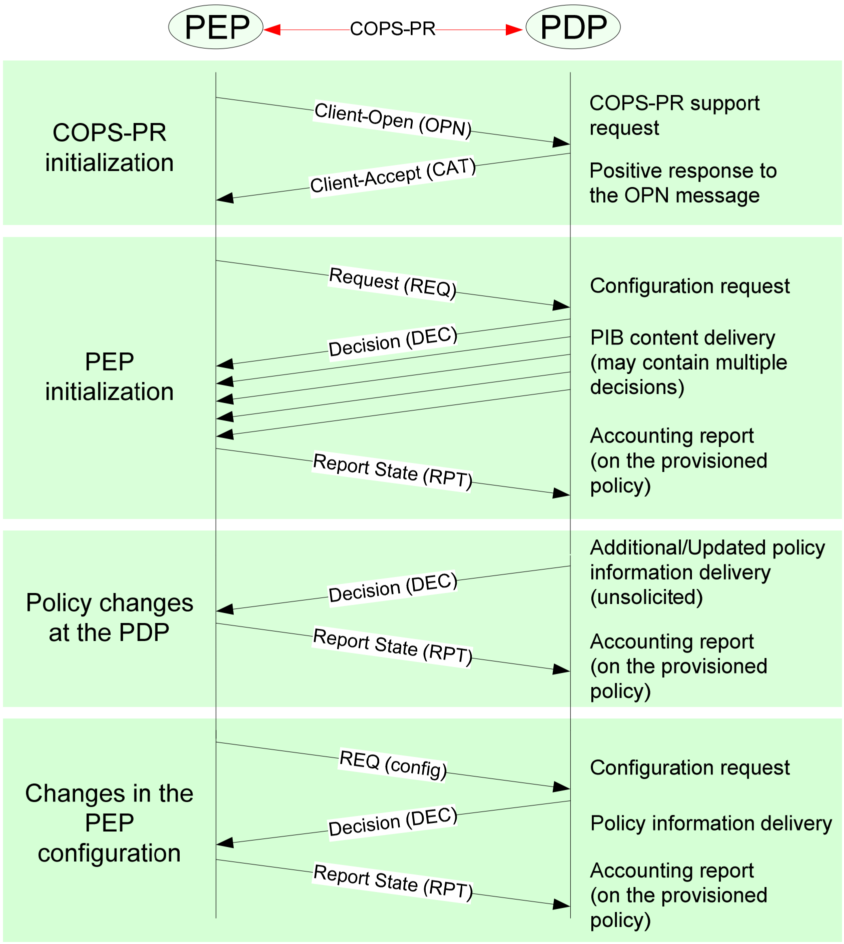

PEP
COPS-PR
PDP
COPS-PR initialization
Client-Open (OPN)
Client-Accept (CAT)
COPS-PR support request
Positive response to the OPN message
PEP initialization
Request (REQ)
Decision (DEC)
Report State (RPT)
Configuration request
PIB content delivery (may contain multiple decisions)
Accounting report (on the provisioned policy)
Policy changes at the PDP
Decision (DEC)
Report State (RPT)
Additional/Updated policy information delivery (unsolicited)
Accounting report (on the provisioned policy)
Changes in the PEP configuration
REQ (config)
Decision (DEC)
Report State (RPT)
Configuration request
Policy information delivery
Accounting report (on the provisioned policy)

Figure 23: COPS-PR signalling betwen PDP and PEP

There are other intra-domain protocols for network policy delivery, such as the one developed by the Configuration Management with SNMP Working Group (SNMPConf WG) [100], in charge of mapping the Policy WG framework to SNMP [21], defining the Policy Based Management MIB [213]. The model COPS-PR/PIB has some advantages over the SNMP/MIB. The fact that it is stateful enables it to maintain synchronization between the PDP's PIB and the PIB installed in the PEP, constantly checking its validness. Moreover, COPS avoids the problem of multiple management stations of SNMP and the concurrent access to the MIBs given that the PDP has exclusive access to its PEPs. But the most significant improvement is the efficiency gain. The COPS protocol reduces the number of messages exchanged and its complexity, allowing large messages and a greater granular access to the PIB.

A very usual way for transmitting the policies from PDP to PEP is using either the Telnet protocol or the SSH protocol. Table 7 presents the intra-domain communications protocol comparison where it can be seen how these protocols are very poor in optimization terms. This is because with these two protocols it is needed to send all the configuration commands to each of the devices and these may vary depending on the device. Therefore, they use too much bandwidth, offer much less flexibility and the policy delivery cost is also much higher.

	Telnet	SSH	SNMP	COPS-PR
Flexibility	very low	very low	high	high
Funcionality	very high	very high	medium	high
Fiability	high	high	medium	very high
BW use	very high	very high	very low	low
HW requeriments	very low	very low	high	low
Security	very low	very high	high	high
PBNM	no	no	yes	yes
Policy delivery cost	very high	very high	medium	medium

Table 7: Intra-domain communication protocol comparison

3.3.2. Inter-domain communication protocol

An inter-domain communication protocol is needed in order to communicate two peer bandwidth brokers managing DiffServ domains. The idea of this first implementation is to keep it as simple as possible, but with the minimum features of end-to-end inter-domain communication for QoS managing and IPv6 support. The simplest protocol for bandwidth brokers' interrelation is the Simple Inter-domain Bandwidth Broker Signaling (SIBBS) protocol, even though it is not standardized. Furthermore in [148] the SIBBS protocol is proposed as a signalling protocol when scalability is not a factor. Some other proposals have been made for inter-domain communication protocols but will be discussed in later sections.

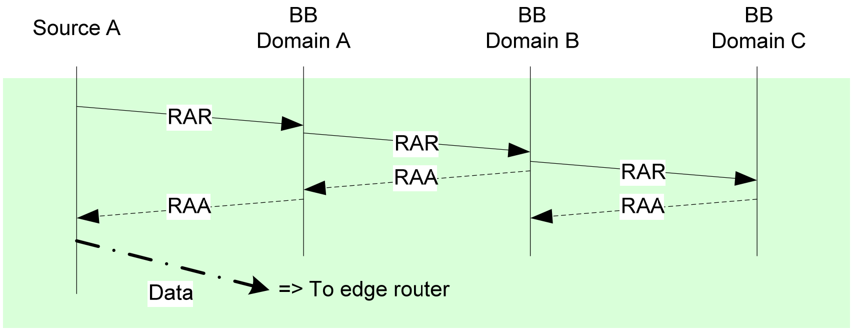


Figure 24: SIBBS signaling between three domains

The Internet2 project [102] developed the SIBBS protocol, the first inter-domain signaling protocol for this kind of architectures. SIBBS is a very simple TCP based protocol to be used between bandwidth brokers managing different domains. A Resource Allocation Request (RAR) message is sent by a bandwidth broker to its peer with information related to the QoS request and the other parameters of the service. A Resource Allocation Answer (RAA) message containing the answer to the RAR is sent back by the other BB peer. If the required resources are available, the request is propagated recursively through the inter-domain path to the last bandwidth broker. This last bandwidth broker returns a RAA message to its immediate downstream bandwidth broker and the process is continued until it reaches the original bandwidth broker. This is concluded with an admission of the QoS request. However, we must highlight the fact that SIBBS may accept data flow even if this flow is later rejected in another domain along the domains' path, as can be seen in Figure 24. This behaviour might imply the loss of prematurely sent packets and, at minimum, implies that these packets will not have applied the required QoS.

3.4. State of the art

The first organism to define the basic architecture of a bandwidth broker [151] was the Internet2 QBone Bandwidth Broker Advisory Council (I2-QBBAC) in the year 2000. This group stated that a bandwidth broker must be basically composed of by an intra-domain communication interface (it proposes the use of specific network policy delivery protocol such as COPS), an inter-domain communication interface for the transmission of the policy decisions between the PDPs of different domains, a database that contains the network topology and by an external or embedded QoS/policy management service which should be based on SLAs that contain the SLSs and RARs. They developed the SIBBS protocol. This architecture can be seen in Figure 25.

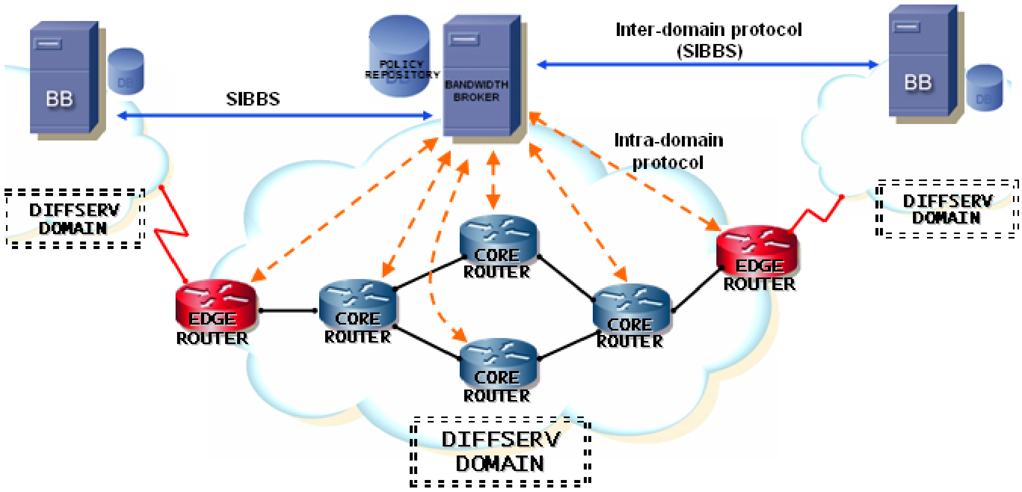


Figure 25: Bandwidth Broker architecture defined by I2-QBBAC

Therefore, according to the aforementioned work of the DiffServ, Policy and RAP WGs and the architecture proposed by the I2-QBBAC, the implementation of a bandwidth broker for the centralized management of DiffServ domains should be composed by a policy delivery intra-domain protocol (COPS-PR) and an inter-domain interface (SIBBS), a database containing the network topology and the policies, and finally a QoS policy management server which should be based in SLAs and SLSs. The bandwidth broker has to manage the kinds of DiffServ PEPs: the domain access routers (edge routers) and the core routers. The SLA and SLS may be configured statically at the PDP by an administrator and the service of policy delivery is then realized by the bandwidth broker receiving a resource allocation request and configuring the routers at the edges of its domain with the set of parameters for the PHBs and the Traffic Conditioning mechanisms.

Even though Internet2 project no longer considered the Bandwidth Broker architecture to be applicable for this specific network in 2001, due to the very high speed acquired by this testing network, the I2-QBBAC specifications have helped as a base for many posterior implementations to be used in networks with QoS needs.

The first practical implementation of a bandwidth broker was carried out by the University of Kansas [197] in 1999 and was based on the early I2-QBBAC's definitions. Another sooner implementation was that of UCLA [193] which stands out for being the first to use COPS for the intra-domain communication. As a result of this working group's tasks, the concept of Bandwidth Broker was defined for the first time in a RFC [156]. One of the most outstanding implementations is that developed by UNSW in JAVA language which supports the COPS and COPS-PR protocols in 2003 [126], based on I2-QBBAC, Policy WG and RAP WGs definitions.

Other implementations of bandwidth brokers [41], [76] and [78] did exist in 2006, but they do not use any standard intra-domain protocol nor do they use early versions which have not yet to be standardized. There were also other implementations which, even though they were not formally bandwidth brokers, realize part of their functions or are integrated in a NMS with AAA. These latter implementations were basically oriented to manage QoS's Mobile IPv6 but they either support COPS-PR for exclusively IPv4 networks [164] or they are simulations that support COPS for IPv6 networks but not COPS-PR [136].

It must be noticed that, besides the mentioned bandwidth brokers, there also existed commercial bandwidth brokers. There were a few and it was only publictly known the protocols supported. By December 2006 all of them functioned exclusively for IPv4 networks and most of them were only capable of managing a single domain.

The state of the art show how the existing bandwidth brokers' capabilies have evolved throughout the time being the latest the more complete and standardized one. It can be concluded that none of the analyzed bandwidth brokers support the IPv6 protocol.

3.5. BBv6 design

This section describes the design and the architecture of a Bandwidth Broker system with IPv6 support (BBv6) and it specifies the functions of its modules.

The BBv6 is described as a bandwidth broker able to manage in a centralized manner the DiffServ domain's QoS which is formed by IPv4 and/or IPv6 networks, and can interact with a peer BBv6 managing a neighboring domain. This architecture has been designed according to the work of the DiffServ, Policy and RAP WGs and the I2-QBBAC. The aim of this project is to implement a system as optimized and standardized as possible. Therefore the BBv6 will support the COPS-PR and the SIBSS protocols, both with IPv6 support, and will be able to manage the full DiffServ technology with two different PEP topologies. The main characteristics of this system are:

- Optimized management of intra-domain policies.
- Optimized management of inter-domain policies.
- Storage of policies, topologies and contracts in a database.
- Remote adminitration with a secure protocol.
- Static management of SLAs and SLSs.
- Support for IPv4, IPv6 and dual networks.
- Support for a full DiffSev domain, with edge routers and core routers, and all the needed traffic conditioners mechanisms.
- Open code

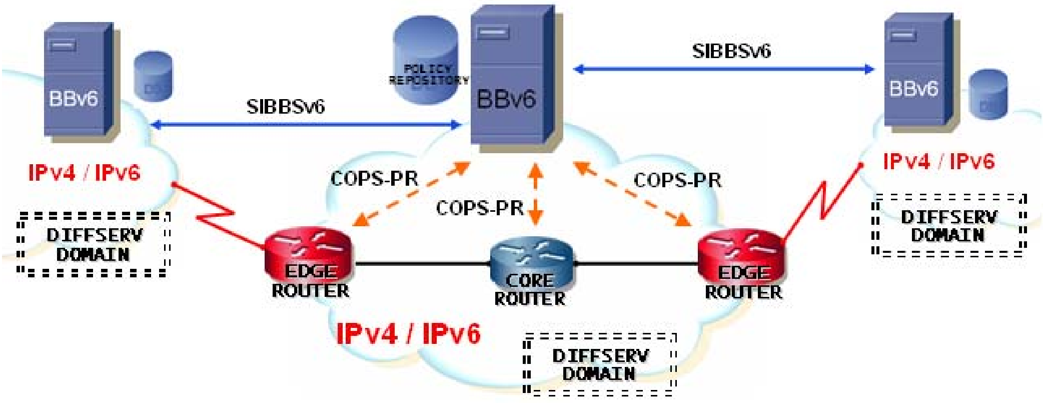


Figure 26: Proposed BBv6 architecture

The system has been designed in modules in order to be easily adaptable to new protocols and technologies. Therefore the system has been divided into six modules, as it can be seen in Figure 27. On the one hand there are the administration modules: the core module, the database module and the management module. On the other hand there are the operational modules: the intra-domain communications module, the inter-domain communications module and the PEP module.

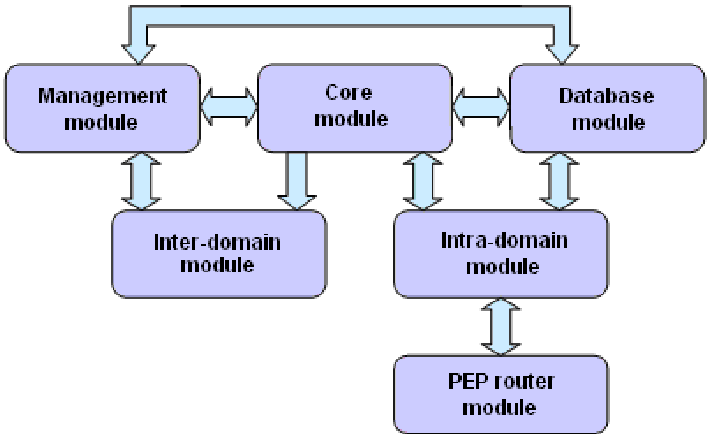


Figure 27: Module diagram of the BBv6 system

3.5.1. Core module

The main module of the BBv6 system is the core module. This module articulates all the QoS control process in the managed domain and interrelates with the intra-domain communications module, the inter-domain communications module, the management module, and the database module.

The aim of this module is to assure the proper functioning of the system and the stability of the system in case of failure of any module. Therefore is responsible for taking the system decisions. It is also in charge of the secured access control.

This module has the responsibility of initialize the rest of the modules.

3.5.2. Management module

The module allows the administrator to manage and remotely configure the whole system. This module interrelates with the core module, the database module, and the inter-domain communications module.

The aim of this module is to allow the SLA and SLS management by a remote administrator.

This module has two parts, the client submodule and the server submodule. The former is used by the remote administrator and it presents a CLI interface. The latter will be responsible for interacting with the rest of the system's modules.The protocol used for the secured connections between administrator and BBv6 is SSH.

The remote administrator has de capacity to create, modify and delete the client's SLAs and SLA's SLSs [12][75]. Within this system, every bandwidth broker's client has a contract or SLA, which controls all the relationship parameters. Inside every SLA there are a series of SLSs which rule the technical parameters (reserved BW, aggregated delay, type of service requested, etc.) of the resource reservation at the domain entrance. Therefore every BBv6 may have many SLAs from different clients and each one may have many SLSs for every client's specific needs. All these policies are stored in the bandwidth broker's database.

The new interface allows the management of both of them and the implementation of the edge router support allows the marking of the packets which satisfy the SLS requirements on entrance of the domain.

3.5.3. Inter-domain module and the SIBBSv6 protocol

The inter-domain communication module allows the interrelation with another peer BBv6 in adjacent domains. Even though this module is initialized by the core module, it will only establish communications the management module.

The aim of this module is to manage efficiently the resource reservation between DiffServ domains managed by BBv6 which are associated. This module is directly related to the management module because it functions statically from the basis of preconfigured SLAs and SLS by the BBv6 administrator of each domain.

Each BBv6 has a unique SLA for each peer BBv6 in its database and the inter-domain communication between BBv6 peers is done on a point-to-point basis. This way a BBv6 is only capable of interacting with its neighbours. To reach further domains the BBv6 will use a predefined BBv6 as a default gateway.

The SIBBS protocol does not support IPV6 and have some other lacks. Therefore it has been modified and the signalling protocol between peer BBv6 will be the new SIBBSv6 protocol.

The SIBBSv6 protocol

This protocol is an enhancement of the SIBBS protocol proposed by I2-QBBAC specially designed for this BBv6 project. SIBBS does not support IPv6and moreover, as mentioned before SIBBS has some lacks in the signalling process. The new protocol SIBBSv6 solves both problems.

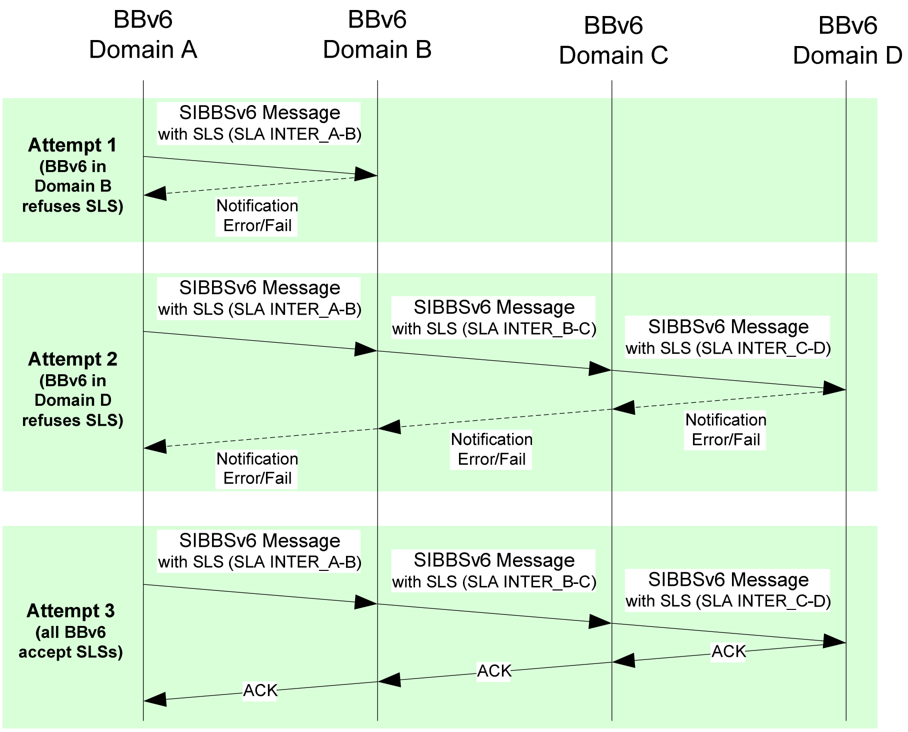


Figure 28: Inter-domain signaling sequence with SIBBSv6

In the SIBBSv6 protocol the resource reservation between domains is now carried out through SLAs and SLSs. Each BBv6 has a unique SLA for each BBv6 peer in its Database. All the SLSs in this SLA will determine the policies at the entrance of the domain for the traffic incoming from the peer domain. Therefore the BBv6 now becomes the “client” and the SLSs between domains will be associated to the same SLA, even if they are from different external clients. This way we prevent confidential client data from being sent to the Service Provider of the peer domain.

It has been also modified the signaling process so that the source client cannot send the data until all BBv6 along the domain path have accepted it. If a domain can not accept the QoS requiriments (in form of SLS), it sends a Notification Error/Fail message and the process is stopped (Figure 28). This consequently implies a greater delay but avoids sending data which could be not accepted by a remote BBv6 because of the QoS lacks.

Furthermore, the SIBBSv6 protocol has been adapted to support IPv6.

3.5.4. Intra-domain module

The intra-domain communications module allows the DiffServ QoS policy delivery to the nodes under management. This module interrelates with the core module, the database module and the PEP router module.

The aim of this module is to manage the QoS policies delivery as optimized and centralized as possible. The communications protocol is COPS-PR. This protocol needs a PIB for each different PEP and also needs to know which routers are part of the topology under management. Therefore, this module has to communicate with the database module. This module is divided into two parts. The PDP submodule is located in the BBv6 and is the responsible for interacting with the core module and the database module. This submodule has various PIBs located in the database, a generic for all the core routers and one for each edge router. The PEP submodule is located in every router under management which have a local PIB where it stores the own QoS policies delivered by the PDP. This submodule will communicate with the PEP router module which translates and installs the QoS policies from the PIB to the router's configuration.

The PIB structure has been designed according to the DiffServ WG [24] and RAP WG [179] definitions and has been provided with IPv6 support for IPv6 nodes. The previously configured COPS-PR protocol has also been modified in order to support IPv6, mainly the headers and the redirection functionality.

3.5.5. PEP Router module

This module is installed in all the domain's nodes under management. All this nodes have support for IPv6, DiffServ, COPS-PR and have a GNU/Linux operative system. It only interrelates with the intra-domain communications module, concretely with the PEP submodule.

The aim of this module is to translate the QoS policies existing in the PEP router's PIB of the intra-domain communications module, so they are usable by the router's operative system and implement them; thus they are effectively configured and the router is able to provide the contracted QoS to the data traffic that traverses it. If the operative system of a router changes, the only module in the BBv6 system which has to change is the PEP router module, as it is the only one specific for each operative system.

This module is able to support all the Traffic Conditioning mechanisms described in the DiffServ architecture [12]: classification, metering, shaping, policing (dropping) and marking of packets, in addition to queuing. Hence, the new PEP module is ready to manage all the PHB traffic types: Expedited Forwarding (EF) [35] for delay sensitive premium traffic, Assured Forwarding (AF) [80] for non-critical priority traffic and BE for the rest of the traffic.

3.5.6. Database module

This module is in charge of managing the database access. It does interrelate with the core module, the management module and the intra-domain communications module.

The aim of this module is to manage the communications with an open source data base which supports IPv6 connections.

This database stores the high level QoS policies in form of SLAs and SLSs, created by the administrator in the management module, the peer BBv6 addresses, provided by the inter-domain communications module through the management module, and the domain's logical topology, provided by the intra-domain communications module.

3.6. BBv6 implementation

This section describes the implementation of the BBv6 system and its modules.

This BBv6 implementation has been based on the embryonic code of a bandwidth broker.

This implementation has been provided by the University of New South Wales (UNSW), Australia, who has authorized expressly URL to modify their final 2003 version [126].

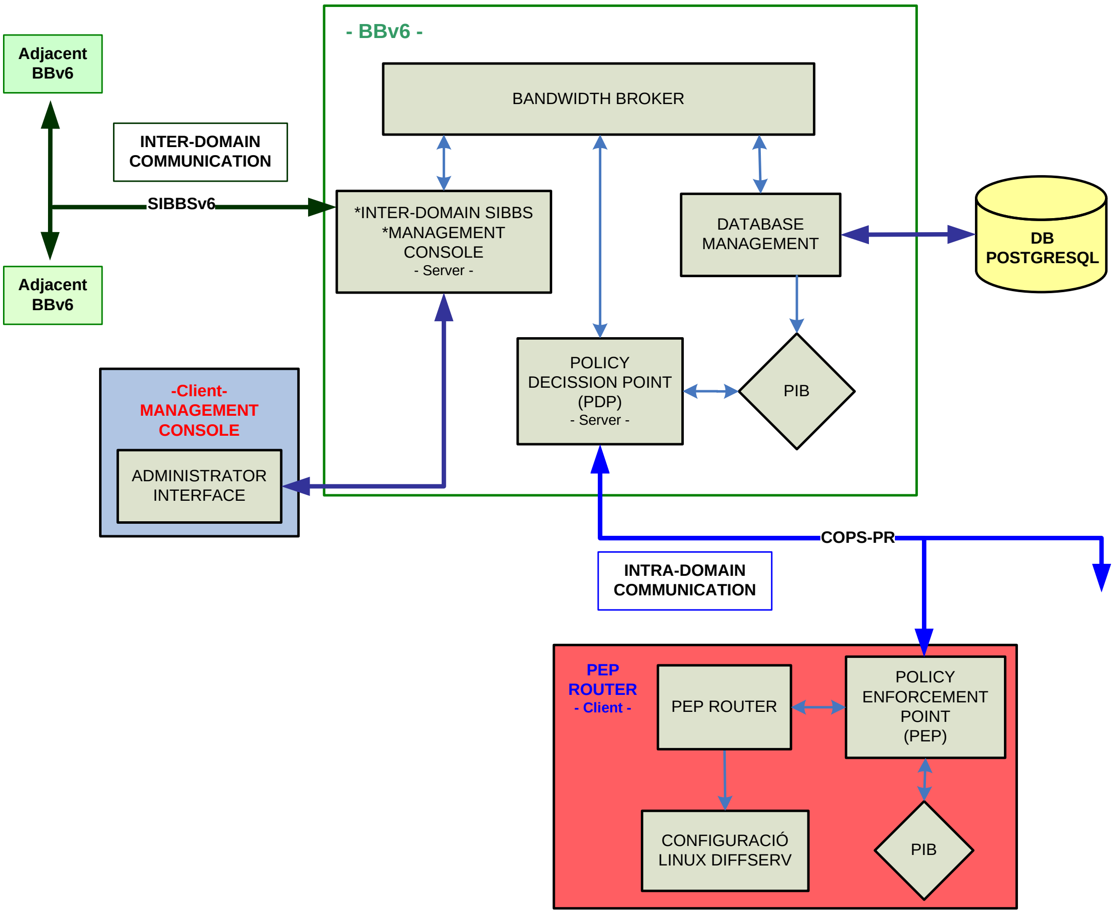

- BBv6 -
BANDWIDTH BROKER
Adjacent BBv6
INTER-DOMAIN COMMUNICATION
SIBBSv6
*INTER-DOMAIN SIBBS *MANAGEMENT CONSOLE - Server -
DATABASE MANAGEMENT
DB POSTGRESQL
Adjacent BBv6
POLICY DECISSION POINT (PDP) - Server -
PIB
-Client- MANAGEMENT CONSOLE
ADMINISTRATOR INTERFACE
COPS-PR
INTRA-DOMAIN COMMUNICATION
PEP ROUTER - Client -
PEP ROUTER
POLICY ENFORCEMENT POINT (PEP)
CONFIGURACIÓ LINUX DIFFSERV
PIB

Figure 29: Block diagram of the modules and submodules of the BBv6 architecture

The original code, developed in JAVA language and open coded, supports COPS and COPS-PR and it is based on the definitions of the Policy and RAP WGs and the I2-QBBAC. The BBv6 improves the code by implementing some of the deficits detected. The COPS-PR protocol implemented does not fit the standard as it does not have PIB and it is only capable of managing core routers with very limited DiffServ QoS capacities. Similarly the SLA and SLS management is based in the non-standarized definitions of the I2-QBBAC. The original code only supports IPv4.

The new implementation has transformed the original bandwidth broker into a dual stack device which implements all the requirements to support IPv6 in all the modules. The COPS-PR protocol and the DiffServ technology are now supported completely. The database and the management console have been reviewed and modified in order to standardize the SLA and SLS management according to IETF specifications. Furthermore it has been fully modified an early implementation of the SIBBS protocol and it has been enhanced to the SIBBSv6 protocol. Details of the JAVA classes of the BBv6 architecture are detailed in Appendix B.

3.6.1. Core module

The core module is the main part of the system. This module initializes all the other modules located in the BBv6. Figure 29 illustrates the block diagram of the architecture.

It contains the control plane of the intra-domain communications module's PDP in order to decide how to proceed with every incoming COPS-PR message.

The module executes two JAVA threads. The first thread waits, manages and authenticates the remote console request and the remote peer BBv6 requests coming from other domains. The second thread reviews the SLS status against the database module every certain period.

3.6.2. Management module

The management module implements the communication protocol for the SLA and SLS creation between the administrator and the BBv6 (Figure 30).

The part of the management module located in the BBv6 is in charge of managing the messages coming from the remote administration and makes the proper actions against the database module in order to save the new configurations. It is also responsible for managing the messages against the inter-domain communications module if the client needs end-to-end QoS and destination is in another domain.

The part of the remote management module manages communications and authentication with the BBv6. It provides a CLI interface for the configuration of QoS policies in the BBv6.

The fact that the bandwidth broker of the UNSW only permits core router configuration, implies that it is not able to configure a RAR on entering the domain, in accordance with the I2-QBBAC specifications which it is based on, even though its management console indicates the capacity to carry out such configurations. The management console has been completely redesigned substituting the RAR concept with the standardized concept of SLA/SLSs [12][75].

3.6.3. Inter-domain module

The inter-domain module is responsible for managing the QoS resource request coming related to other peer domains. Every two domains will have a contract or SLA, with different SLSs, so that there will be statically and proactively assigned traffic resources for the traffic between both domains. The SIBBSv6 protocol manages the QoS resources request from a domain towards another domain. This other domain may be located several domains away as long as all the domains along the path have a BBv6 supporting SIBBSv6.

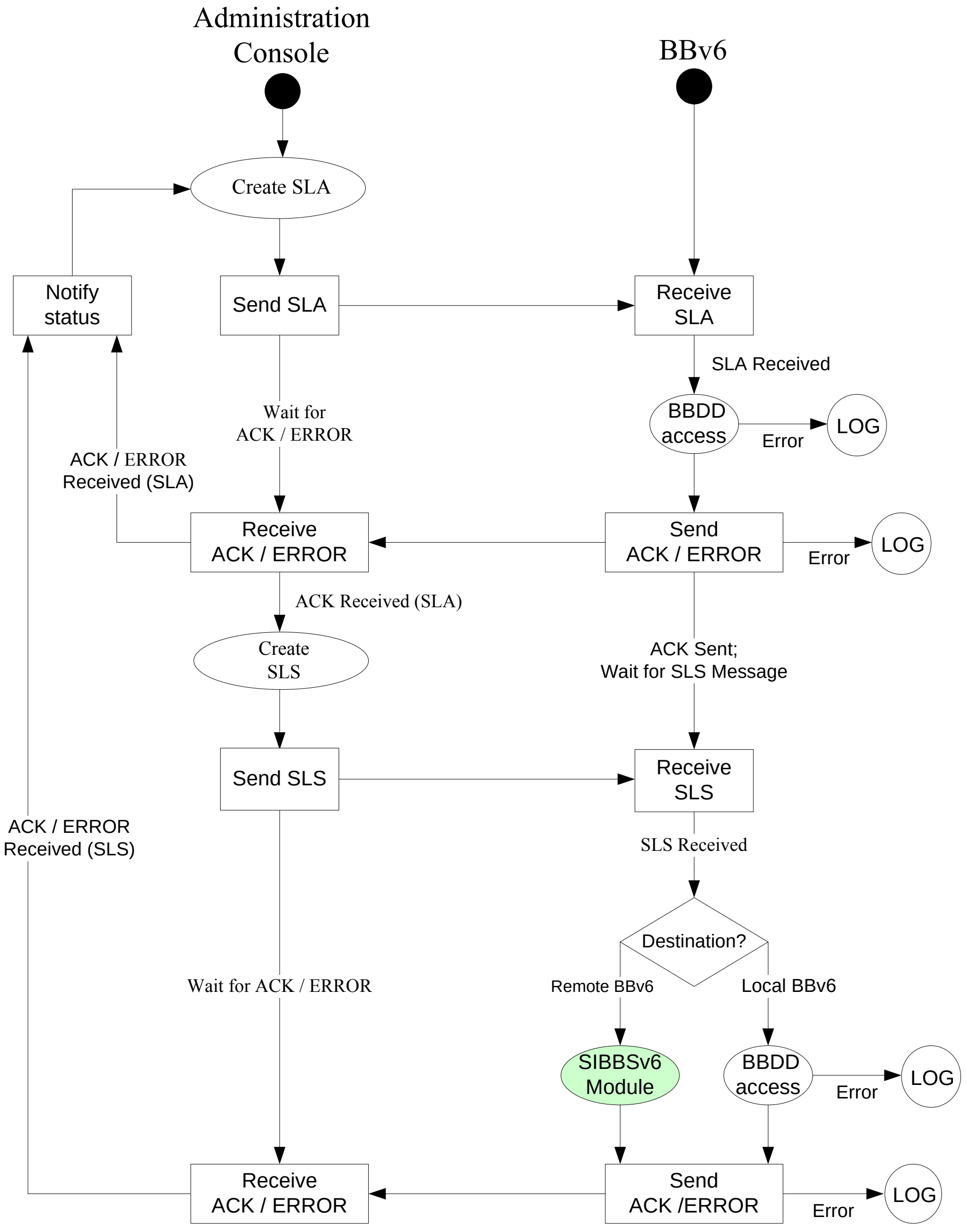


Figure 30: Flow diagram of the communications between administrator and BBv6

When a remote administrator introduces a new SLS request, and the traffic destination is located in another domain, the resource request against another BBv6 is done through the SIBBSv6 protocol (SIBBSv6/inter-domain module). Therefore, the request comes from the management module as seen in Figure 30. If the traffic destination is not in an adjacent domain, then this peer domain will act as a remote administrator and will request resources to a third BBv6, as it can be seen in Figure 31. This procedure can be repeated cyclically with as many BBv6 as domains in the data path.

Transit BBv6 will wait until the notification of last domain in the data path in order to reserve resources inside its own domain. This was shown in Figure 28. On the other hand, the core module has a Thread responsible of activating the inter-domain module when a SIBBSv6 request is received from another domain.

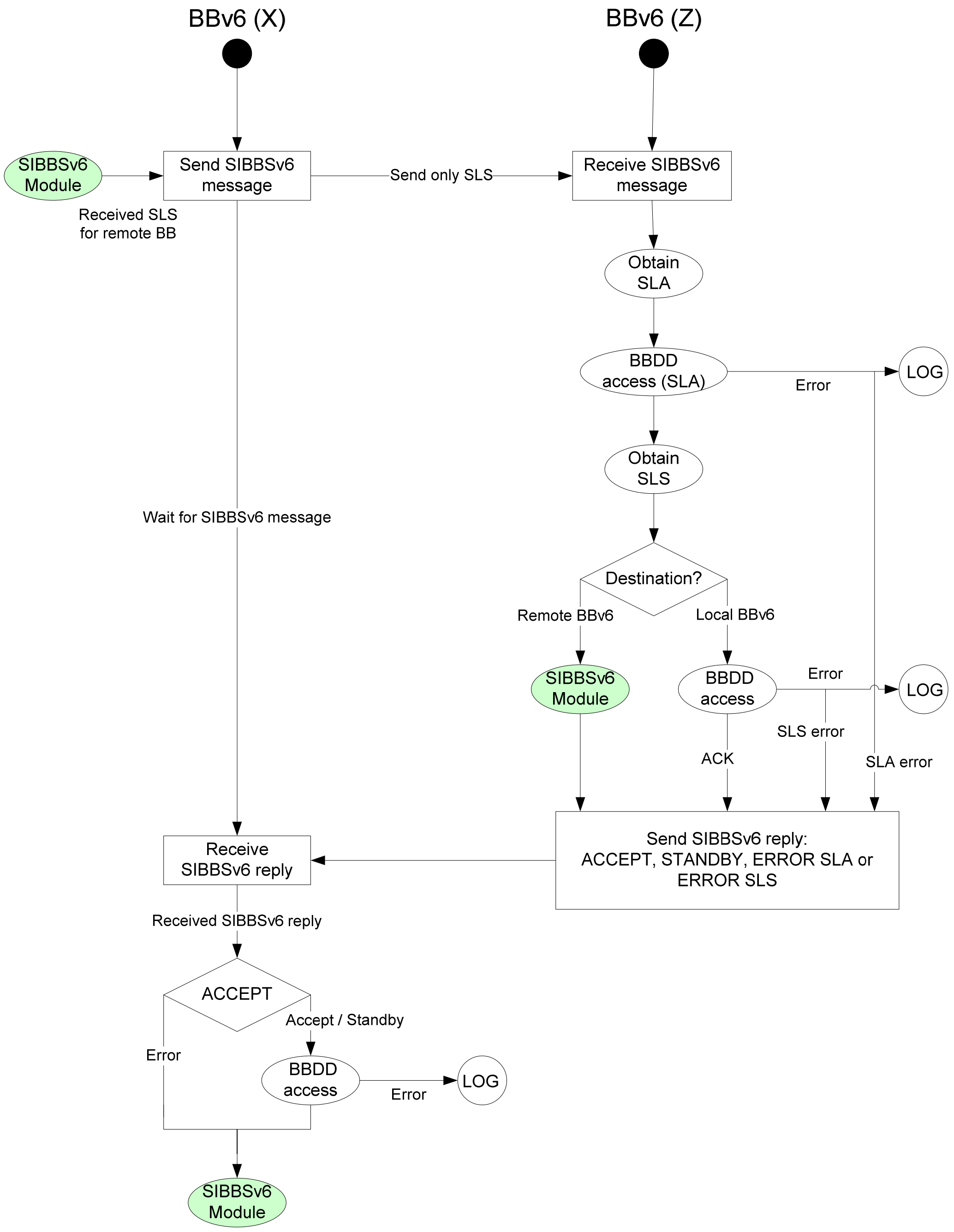


Figure 31: Flow diagram of the SIBBSv6 protocol communications between peer BBv6

3.6.4. Intra-domain module

The protocol used for the intra-domain communication is COPS-PR. The original code from is lacking in several features related to the implementation of the COPS-PR standard that we have corrected, such as the use of the protocol without a defined PIB structure, even though it has the oriented-structure to objects for its creation. Furthermore, does not have implemented the keep-alive functions, redirection in fail case and synchronization for the PDP-PEP disconnection case, all of which reduce the BB capacities to detect and getting over failures. In addition to fully support of IPv6 protocol, it has been implemented the whole protocol, including the keep-alive function, the synchronization function for the PDP-PEP disconnection case and the PEP-redirect function in fail case. This latter function also supporting PDP redirection with IPv6 addresses.

Support for the PIB defined by the DiffServ WG [24] has also been included in addition to the previously determined RAP WG PIB definition [179], which also supports IPv6. The new implemented PIB completely fulfils them.

Therefore the BBv6 architecture completely implements the COPS-PR protocol with IPv6 support.

3.6.5. PEP Router module

The original UNSW implementation has a very simple configuration of this module. It only incorporates support for managing core routers which are able to exclusively prioritize Expedited Forwarding (EF) PHB traffic ahead of Best Effort (BE) traffic. A sit can be seen in Figure 32 it uses DSMARK, PRIO i PFIFO queing disciplines.

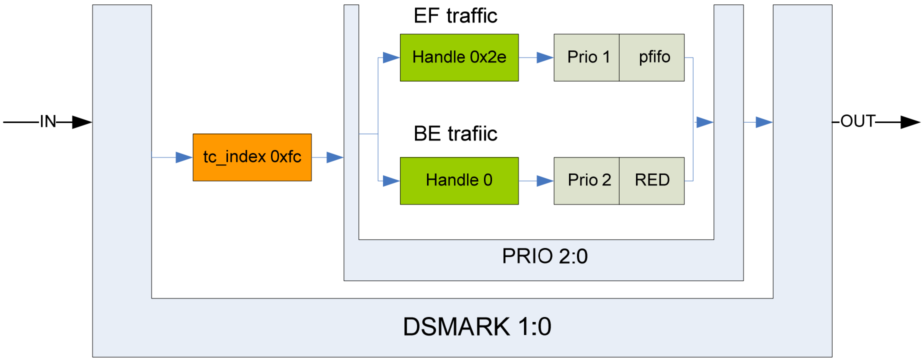


Figure 32: Original queuing system of the core routers by the UNSW

The PEP module installed in every IPv6 DiffServ node is responsible for the configuration of the policies in the PEP's PIB in a computer running GNU/Linux as a router. In the IPv6 DiffServ networks of the implemented testbed the edge routers are able to support all the Traffic Conditioning mechanisms described in the DiffServ architecture in addition to queuing and the core routers are able to manage marked traffic on the basis of the Behavior Aggregate classifier.

The queuing systems used in these GNU/Linux nodes are mainly the DSMARK and Hierarchical Token Bucket (HTB) queuing disciplines as can be seen in Figure 33, even though PFIFO and virtual queues (for the AFxx sub-classification inside the GRED system) have been also used. HTB permits the definition of relative bandwidths for each class of traffic and accepted bandwidth to borrow from other classes.The GNU/Linux kernel has been configured for IPv6 support with all the QoS functionalities available to be used in IPv4 as well as IPv6.

The mapping of the HTB queuing discipline inside the COPS-PR's PIB in the intradomain communication protocol has been implemented through the concatenation of PRC dsSchedulers.

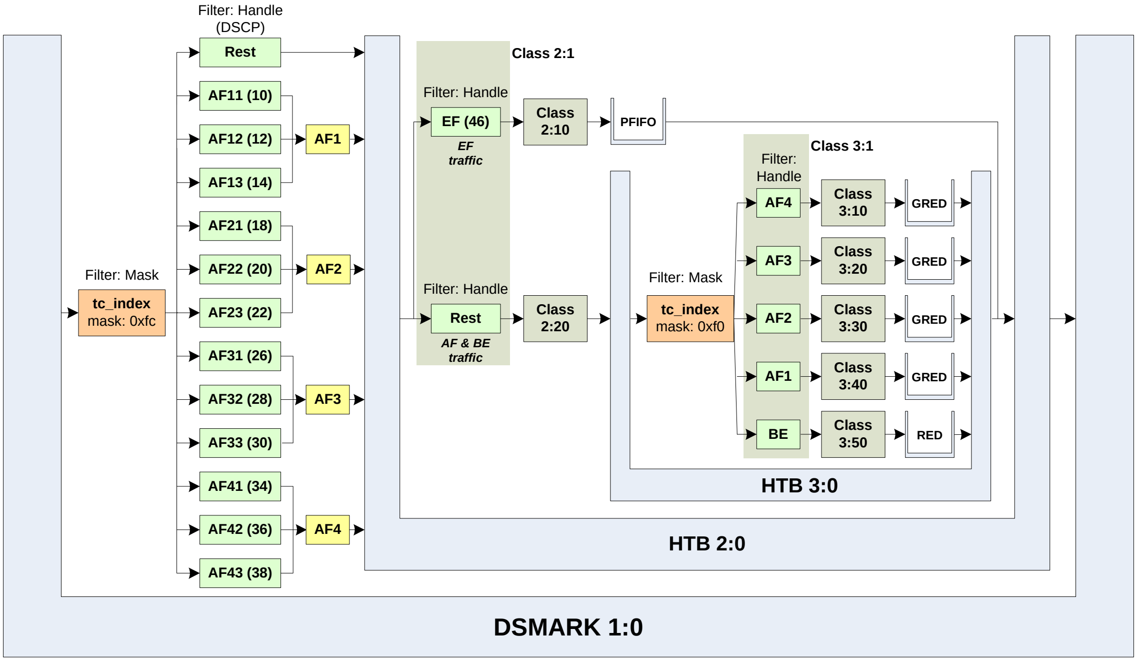


Figure 33: New queuing system of the core routers with GNU/Linux OS

3.6.6. Database module

The Database has been implemented using PostgreSQL v.7.4.6. because its JAVA connector supports IPv6 connections while the original implementation by UNSW used MySQL which JAVA connector did not support IPv6 connections.

The new database uses SQL to store the domains policies such as SLAs, SLSs, the network topology and the peer BBv6 addresses. The SLSs are translated to the PIB for their distribution through COPS-PR to the PEPs, as well as core and edge routers, pointed out by the network topology. Therefore, the relation of routers of the domain allows the installation of the network policies to the routers.

3.7. Evaluation and results

In order to evaluate the proper functioning of the overall BBv6 system, two different tipologies of tests have been carried out. The first tests have been done to evaluate the BBv6 system and the intra-domain communication protocol. The second part of the tests has been done to evaluate the overall system including SIBBSv6, the inter-domain communication protocol.

3.7.1. Evaluation of the intra-domain communication

The performance of the policy-based system has been evaluated through the implementation of a testbed, as can be seen in Figure 34. All the computers of the IPv6 testbed have been upgraded to Intel Pentium IV, 2.8 GHz running Debian GNU/Linux 3.1 with rebuilt Kernel

2.6.9 with all the QoS options. The PEP modules have been installed in the two edge routers and in the core router and the link which is going to be congested is implemented with a 10 Mbit/s Ethernet hub in the network 2001:720:818:4003:2:: /80 with an average performance of 70%.

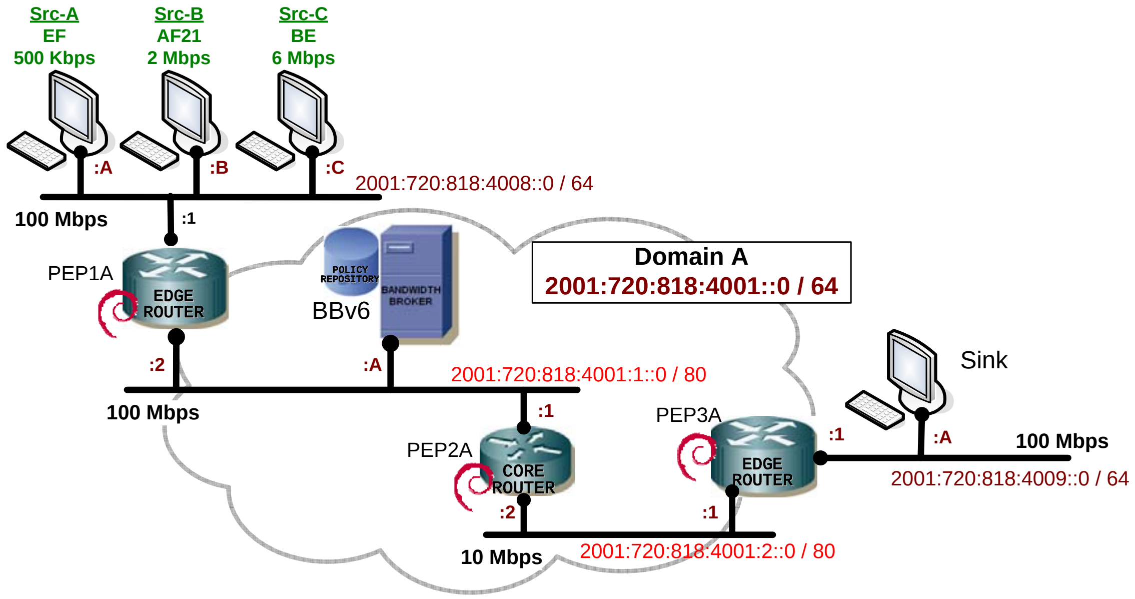


Figure 34: Testbed implemented to evaluate the intra-domain communication

Three UDP traffic sources have been configured to make the tests, which have been generated with Iperf 2.0.2: An EF traffic source which generates 500 Kbit/s and will be marked as EF at the edge router PEP1A, an AF traffic source which generates 2000 Kbit/s and will be marked as AF21 at the edge router PEP1A and a final BE traffic source which generates 6000 Kbit/s. It must be pointed out that with Iperf the datagram size needs to be lowered when using IPv6 addressing to 1450 bytes or less to avoid fragmentation.

QoS policy definition

In order to implement the traffic conditioner required by DiffServ, the PIB in Figure 35 will be installed in the edge router PEP1A. For the sake of simplicity it has been omitted the PRID fields as well as those which are not applicable and therefore have a zeroDotZero value. Two classifier elements can be seen in the PIB, one for each installed SLS with its corresponding IPFilter, which map the data identifying the conformance traffic. The QoS policies for the two SLSs are specified in Table 8.

On the other hand, the PDP sends other PIB policies for the core routers (Figure 36). The traffic conditioner, or shaper, has been implemented using queuing and the dsScheduller with dsMinRate parameters.

Once they are installed the routers are able to classify packets according to EF and AF PHBs.

	SLS 1	SLS 2
Class	EF	AF21
DSCP	46	18
Traffic source	Src-A	Src-B
Traffic destination	Sink	Sink
Bandwidth	1 Mbit/s	3 Mbit/s
Burst	50 Kbytes	50 Kbytes
***out-of-profile* action**	Drop	Remark as BE

Table 8: QoS policies of the SLSs in the PEP1A edge router

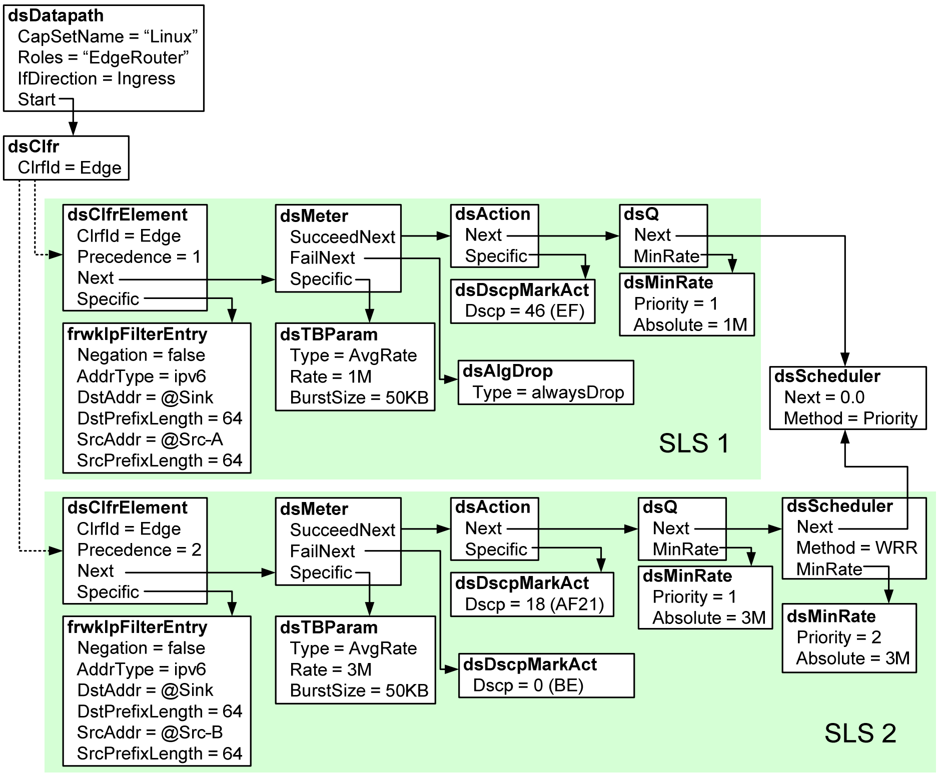


Figure 35: PIB policies in the PEP1A edge router

These policies state that the EF class has the maximum priority in the packet queuing system. The AF classes append the possibility of using GRED to provide discarding probabilities which will depend on the subclass used. The PIB uses the Random Dropper to define GRED.

- EF class: DSCP 46, 10% BW allocated, 100% allowed if no congestion.
- For each AFx class: 20% BW allocated, 90% allowed to borrow from other classes if no congestion. The GRED discarding probability parameters are:
  - AFx1 drop probability: 0.02
  - AFx2 drop probability: 0.04
  - AFx3 drop probability: 0.06
- BE class: DSCP 0, 10% BW allocated, 90% allowed to borrow from other classes if no congestion. TailDrop droping algorithm.

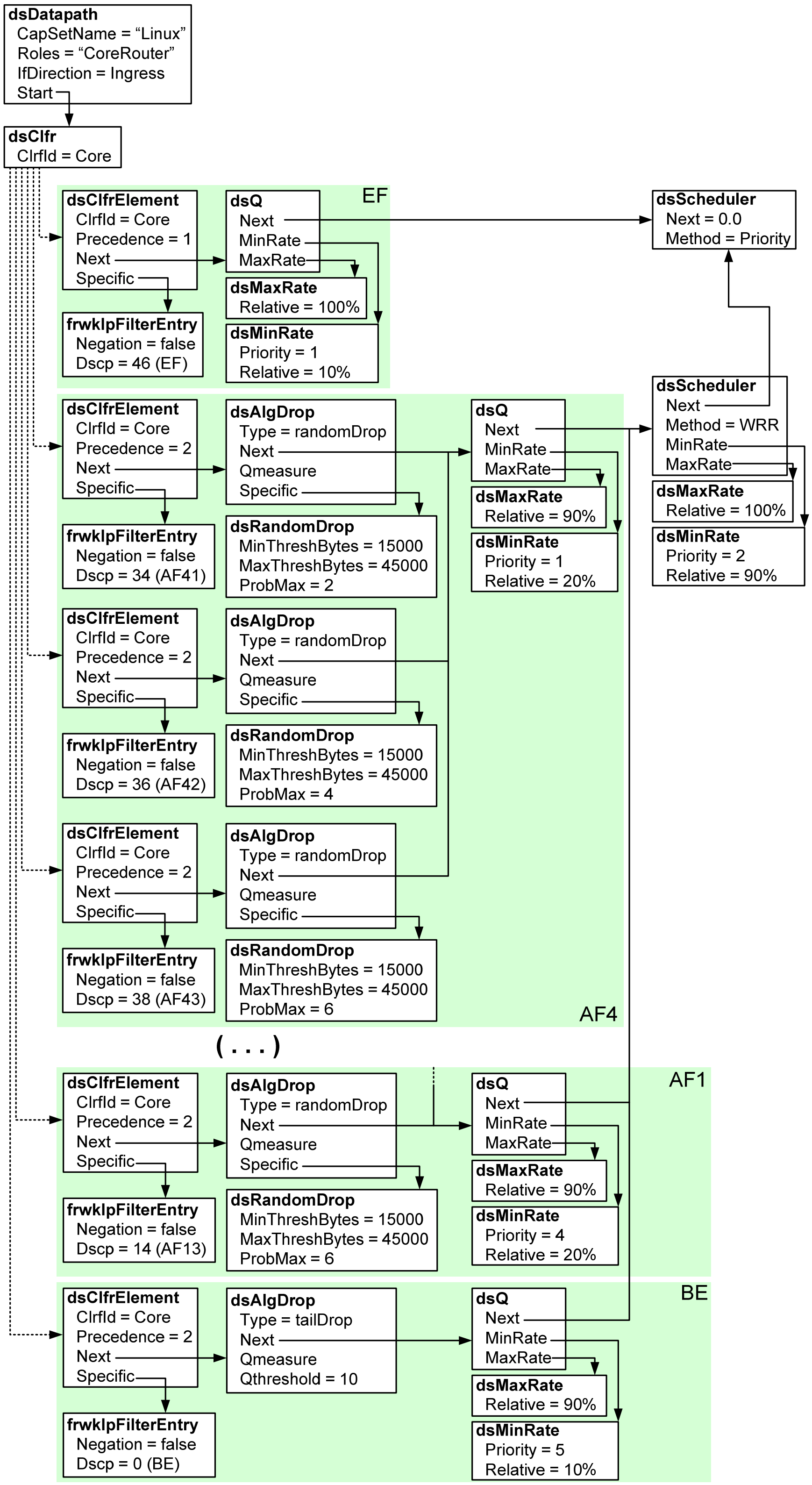


Figure 36: PIB policies in the PEP2A core router

The recommended way for controlling congestion is with random droppers. These dropper are algorithmics and they work indistriminately over any packets in each queue. Traffic separation is done by previous classifiers and meters in form of dsClfrElements and dsMeters. The implemented PIB use the random dropper mechanism to define the behaviour of GRED, which associate the AFx queues.

The mapping of the Hierarchical Token Bucket (HTB) queuing discipline inside the PIB is done through the concatenation of PRC dsSchedulers. In this PIB the AF's and BE's dsScheduller use Weighted Round Robin (WRR) as a service method for their corresponding queue, which serve them using the dsMinRate and dsMaxRate parameters of each queue. The EF's dsScheduller uses strict priority as a queuing service method also using the dsMinRate and dsMaxRate parameters. The dsMinRate parameter fixes the superior threshold and dsMaxRate parameter fixes the inferior threshold.

3.7.2. Results of intra-domain evaluation

Firstly the BBv6 initializes all the nodes in the domain it has stored in its topology database. Once the QoS policies are set, the BBv6 installs the policies of the BA classifier from the PIB at the core router PEP2A. Then the BBv6 applies the configured SLSs to the edge routers who will apply the QoS policies to incoming traffic.

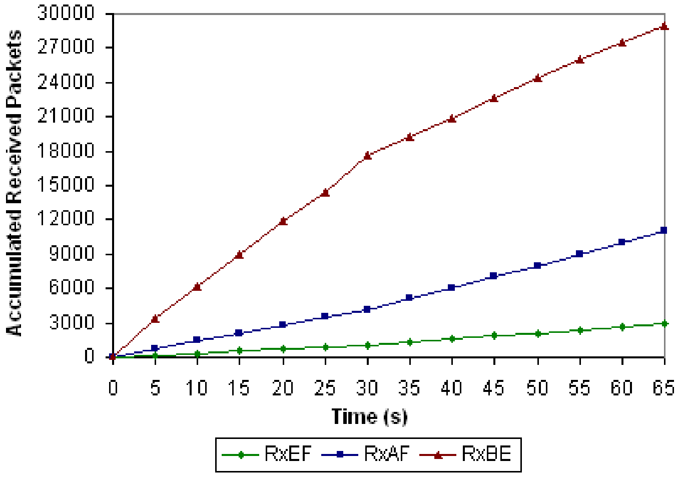


Figure 37: Packets received before and after the policies are applied by the BBv6

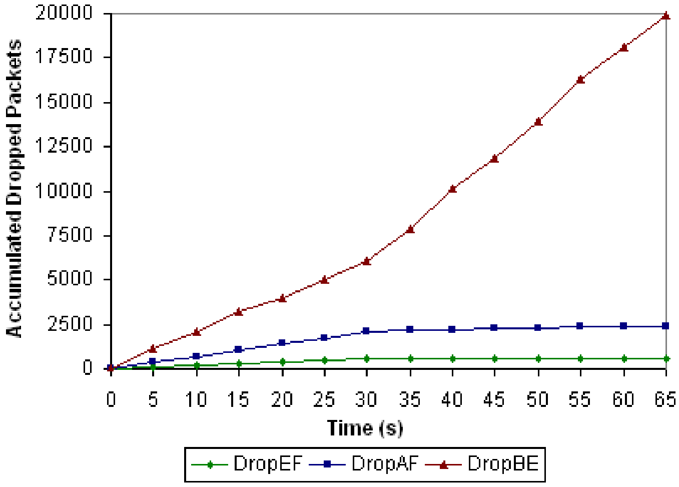


Figure 38: Packets dropped before and after the policies are applied by the BBv6

Figures 37 and 38 illustrate the network behavior initially when the edge routers do not have any policy configured and when these policies are applied in the BBv6 after 30 seconds. The BBv6 then distributes them to the edge routers for the incoming packet treatment. Just at this moment the EF traffic and the AF traffic reduce drastically their loses as it can be seen. From then on, most discarded packets are from the BE traffic source. The figures show the accumulated received packets and the accumulated dropped packets for each type of traffic.

As it can be seen in Table 9, before the policy is applied the system was dropping 33.99% of the EF packets, 33.15% of the AF packets and 25.64% of BE packets. Once the policy is applied, the drops of the EF packets stops (0%), the drops of AF packets are drastically reduced to 3.6% and the BE packets rise to 43.98%.

	Destination Sink	
	Before policies are applied	After policies are applied
EF	33.99 %	0 %
AF21	33.15 %	3.6 %
BE	25.64 %	43.98%

Table 9: Percentage of dropped packets (intra-domain)

These results are correct according to the policies implemented in the core routers. The policy provided maximum priority for EF packets and a smaller drop probability to AF traffic compared to BE.

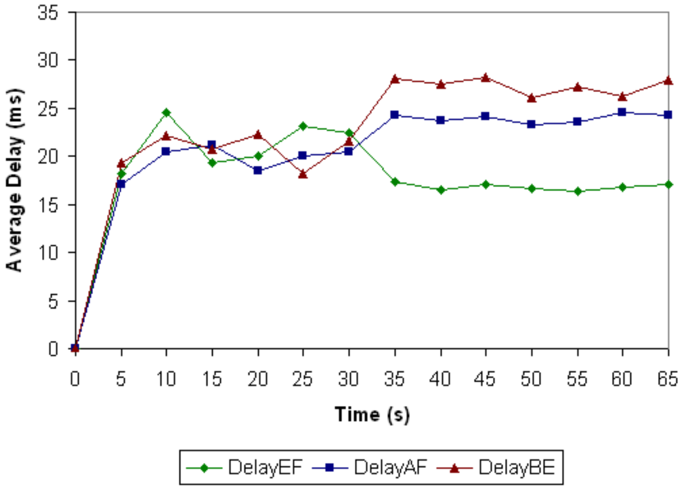


Figure 39: Average packet delay before and after the policies are applied by the BBv6

Figure 39 illustrates the behavior of the packet delay parameter in 5-second intervals averages within the same test where the policies are applied in the BBv6 after 30 seconds. Before policies are applied general packet delay is around 20 ms and all packets behave approximately equal. When policies are applied EF packets delay stabilizes at around 17 ms, AF packets delay at around 24 ms and BE packets delay is always over 27 ms. These results are coherent with the policies because EF packets have maximum priority. Once EF packets are shaped at the first edge router they observe virtually no queuing delay. On the other hand AF and BE packets are queued with statistical priorities and drops. Furthermore, BE packets have minimum priority.

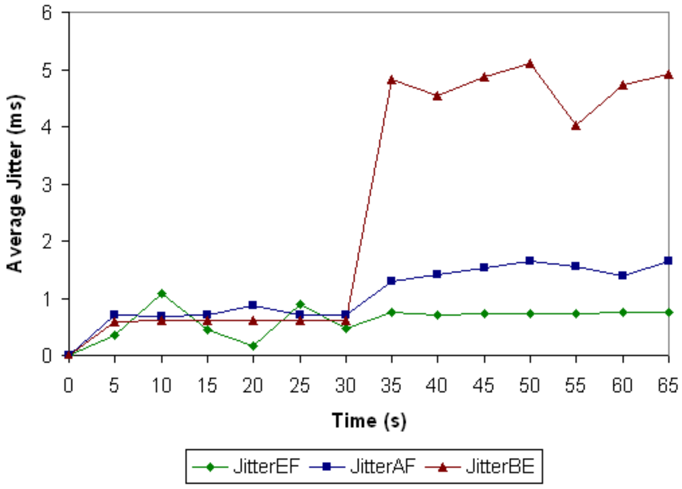


Figure 40: Average packet jitter before and after the policies are applied by the BBv6

Figure 40 illustrates the packet jitter parameter within the same test. The variability in the time delay between packet arrival of each traffic kind as a result of the introduction of queuing can be observed. EF traffic is always prioritized and therefore has minimum jitter. On the other side BE traffic is always statistically conditioned by the rest of the traffic and therefore presents a higher variation.

Even though delay and jitter characteristics indicate a good performance of the QoS, the small size of the testbed makes it difficult to detect great variances in either of theses characteristics. Moreover, only serialization delay, queuing delay and network delay are present as packets are generated with a traffic simulator and no packetization delay or coding delay are present (sourcing delays).

Other tests were conducted under similar conditions and presented very similar results, even when traffic conditions were modified but congestion levels remained the same.

Therefore, the results validated the correct functioning of the BBv6, the policy management including the creation, the storage and the recuperation, the performance of the policy-delivery system, the COPS-PR protocol, the GNU/Linux testbed and the modules implemented in each node [202]. Therefore are presented as a contribution of this thesis.

3.7.3. Evaluation of the inter-domain communication

The performance of the policy-based architecture has been evaluated through the implementation of a testbed with three IPv6 DiffServ domains. This enhanced testbed is the result of cloning three times the intra-domain testbed as it can be seen in Figure 41. This scenario has three domains with IPv6 networks and DiffServ support, each one managed by its own BBv6. All the computers in the testbed are Intel Pentium IV, 2.8 GHz running Debian GNU/Linux 3.1 with rebuilt Kernel 2.6.9 with all the QoS options. The PEP modules have been installed in the two edge routers and in the core router in each of the three domains. The link which is going to be congested is implemented with a 10 Mbit/s Ethernet hub in the network 2001:720:818:4001:2:: /80 (domain A) with an average performance of 70%.

The goals of this inter/intra-domain scenario are to evaluate the correct performance of the overall BBv6 system, the IPv6 DiffServ nodes, the COPS-PR protocol implementation and the

SIBBSv6 protocol over the IPv6 DiffServ domains.

Six different UDP traffic sources have been configured to carry out the tests which emulate real-time traffic fromfrom two different clients, X and Y with two different SLA. Two EF traffic sources which generate 400 Kbit/s and will be marked at the edge router PEP1A, two AF traffic sources which generate 1000 Kbit/s each and will be marked as AF41 and AF21 at the edge router PEP1A and finally two BE traffic sources which generate 3000 Kbit/s each. The traffic has been generated with Iperf 2.0.2.

Two captures of the COPS-PR protocol can be seen in Appendix B.

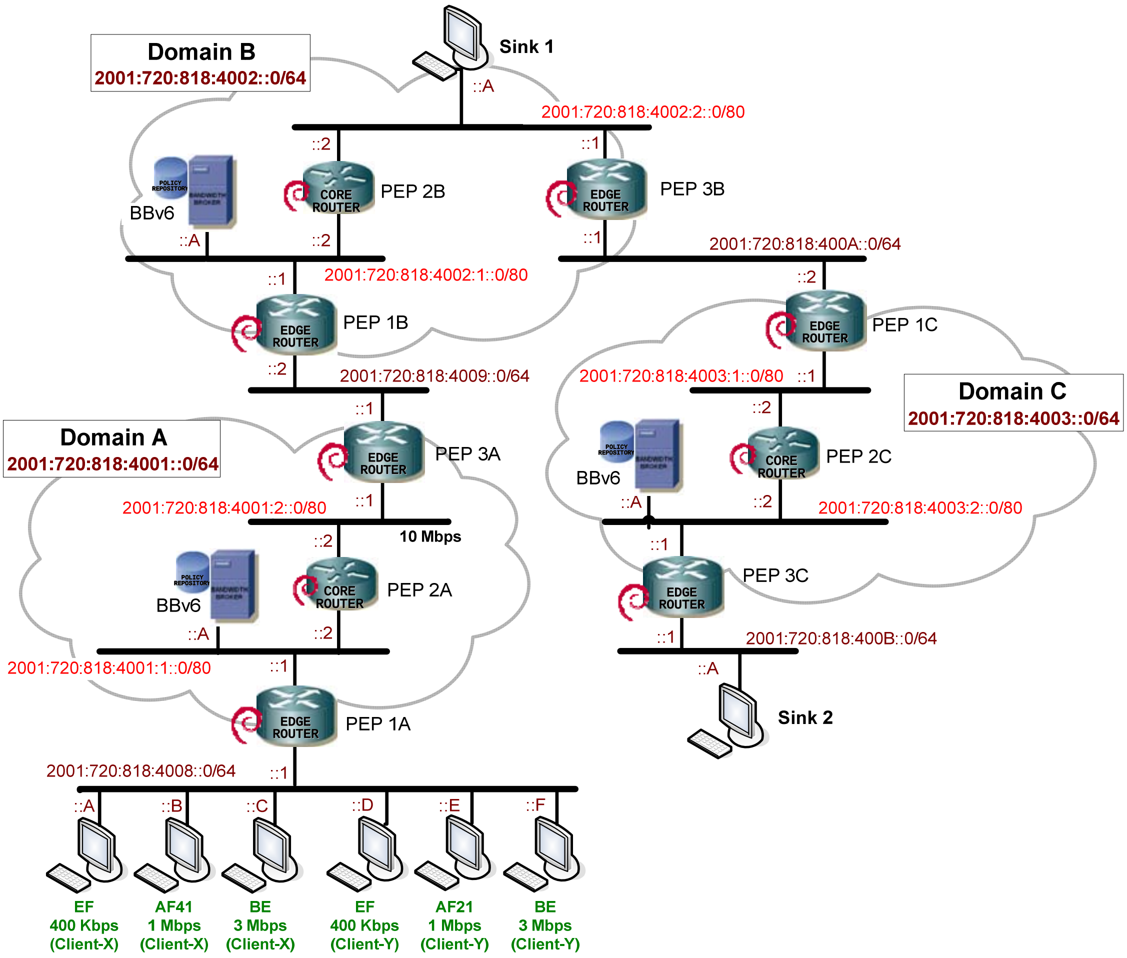


Figure 41: Testbed implemented to evaluate the inter-domain communication

QoS policy definition

The test scenario has three concatenated domains as can be seen in Figure 41. Therefore according to the SIBBSv6 protocol there will be four inter-domain SLAs.

- The BBv6 in Domain A has an inter-domain SLA named INTER_B-A with the SLSs which manage the policies for the incoming traffic from Domain B at the edge router PEP3A (the source of the traffic could also be in Domain C).
- The BBv6 in Domain B has two inter-domain SLAs named INTER_A-B and INTER_C-B with the SLSs which manage the policies for the incoming traffic from Domain A and C respectively at the appropriate edge routers (PEP1B and PEP3B).

- The BBv6 in Domain C has an inter-domain SLA named INTER_B-C with the SLSs which manage the policies for the incoming traffic from Domain B at the edge router PEP1C (the source of the traffic could also be in Domain A).

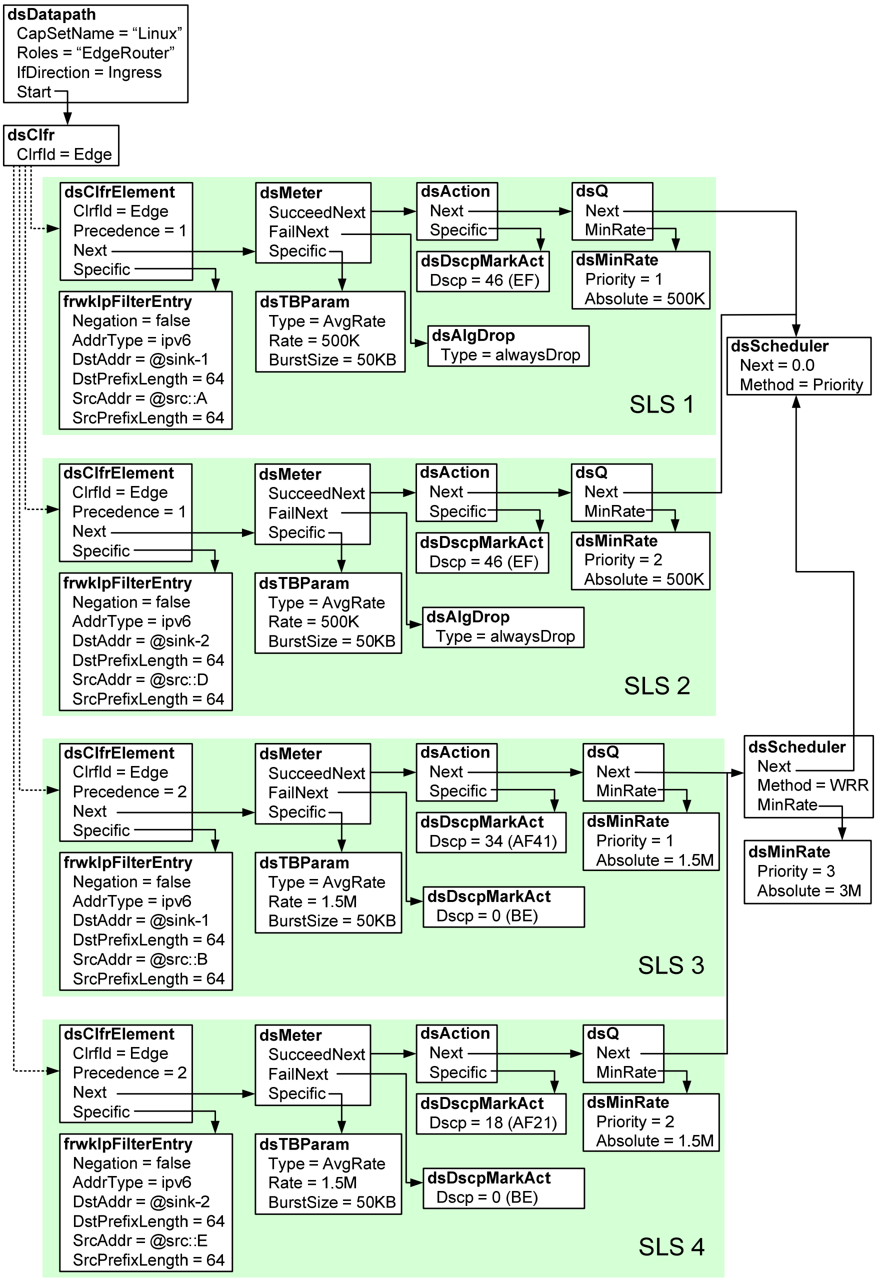


Figure 42: PIB policies in the PEP1A edge router

Besides the inter-domain SLAs, each BBv6 will have their own client SLAs. In the case of the Figure scenario, these will be:

- Client-X has an SLA in the BBv6 in Domain A with two SLSs (1 and 3 in Figure 42) for traffic with destination in Domain B (sink-1). Traffic sources src::A, src::B and src::C are from this client.

- Client-Y has an SLA in the BBv6 in Domain A with two SLSs (2 and 4 in Figure 42) for traffic with destination in Domain C (sink-2). Traffic sources src::D, src::E and src::F are from this client.
- Sink-1 has an SLA in the BBv6 in Domain B.
- Sink-2 has an SLA in the BBv6 in Domain C.

All these policies will be delivered to edge routers by COPS-PR and therefore will have to be mapped in its PIB. The same way it has been done with the PIB of the edge router in the intra-domain scenario, in order to mplement the traffic conditioner required by DiffServ to support the aforementioned SLSs, the PIB in Figure 42 is installed in the in the edge router PEP1A.

The same way, the schema has been also simplified by omitting the PRID fields as well as those which are not applicable and therefore have a zeroDotZero value. There are four classifier elements in the PIB, one for each installed SLS with its corresponding IPFilter, which maps the data by identifying the conformance traffic. All four SLSs have destinations belonging to other domains. The QoS policies for the SLSs are specified in Table 10.

	SLS 1	SLS 2	SLS 3	SLS 4
Class	EF	EF	AF41	AF21
DSCP	46	46	34	18
Traffic source	Src-A	Src-D	Src-B	Src-E
Traffic destination	Sink-1	Sink-2	Sink-1	Sink-2
Bandwidth	500 Kbit/s	500 Kbit/s	1.500 Kbit/s	1.500 Kbit/s
Burst	50 Kbytes	50 Kbytes	50 Kbytes	50 Kbytes
Out-of-profile action	Drop	Drop	Remark as BE	Remark as BE

Table 10: QoS policies of the SLSs in the PEP1A edge router

The policies implemented in core router’s PIBs of the three domains are the same as the one used in the intra-domain scenario (Figure 36). Once the policies are installed, the core routers area ble to classify packets with EF and AF PHBs.

3.7.4. Results of inter-domain evaluation

Firstly, the three BBv6s initialize all the nodes in their domain. Once the QoS policies for the core routers are set in the BBv6s of the domains, the BBv6s install the policies of the core routers using its corresponding PIB. Then, when the administrator configures the new four SLSs in DomainA the BBv6 of this domain, it checks if it has enough resources available to serve each SLS locally. In case of error or in case of denegation of resources, the SLS is placed in Standby mode and a report is sent to the administrator.

In this concrete test, the destinations of the four SLSs are localized in another Domain so BBv6/DomainA has to send SIBBSv6 requests to BBv6/DomainB. The latter BBv6 checks if it can serve the SLSs with the assigned resources to SLA INTER_A-B and with the remaining resources in DomainB. In case of error in any of the SLSs, it will deny the request of the concrete SLS by sending an error/fail notification to BBv6/DomainA which will place the SLS in Standby mode and the administrator will be informed.

The destinations of two of the SLSs are localized in Domain C so BBv6/DomainB sends SIBBSv6 requests to BBv6/DomainC. Then, this BBv6 checks if it can serve the SLSs with the assigned resources to SLA INTER_B-C and with the remaining resources in DomainC. In case of error in any of the SLSs, it will deny the request of the specified SLS by sending an error/fail notification back to BBv6/DomainB which will send an error/fail notification to BBv6/DomainA. BBv6/DomainB will erase the local request but BBv6/ DomainA will place the SLS in Standby mode and the administrator will be informed.

When all the BBv6 along the path have accepted the SLSs, BBv6/DomainC reserves resources in the corresponding SLSs in the SLA INTER_B-C and sends an acknowledgement to BBv6/DomainB. This latter does the same with the SLA INTER_A-B and, finally BBv6/DomainA, on receiving the acknowledgement from BBv6/DomainB, places them in Active mode.

Once the policies have been accepted in the BBv6s of the different domains, the BBv6s will apply the configured SLSs to the edge routers. As stated before, the source client cannot send the data until all BBv6 along the domain path have accepted the related SLS and the policies have been updated at the edge routers.

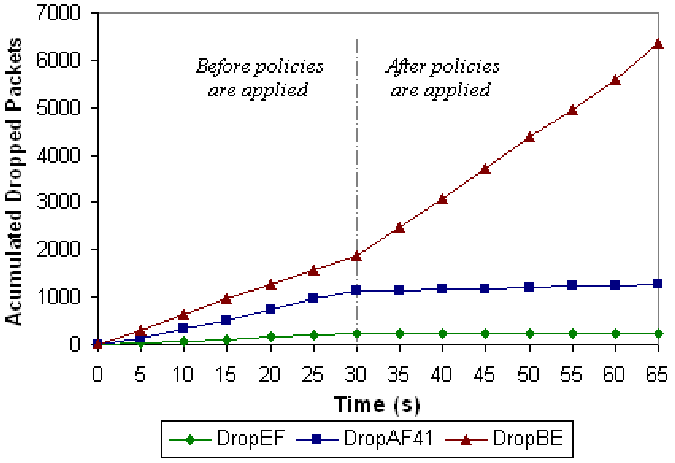


Figure 43: Packets dropped destined to sink-1 before and after policies are applied by the BBv6s

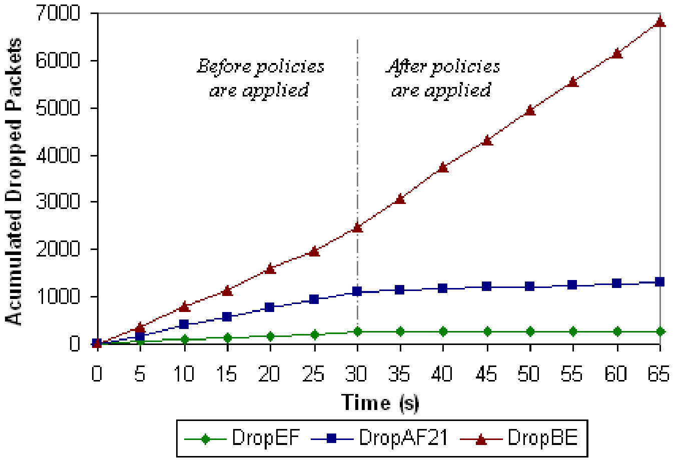


Figure 44: Packets dropped destined to sink-2 before and after policies are applied by the BBv6s

Figures 43 and 44 show the initial behavior of the network before the policies have been configured at the edge routers and when these policies are applied in the three BBv6 and distributed to the edge routers for the incoming packet treatment after 30 seconds. The figures show the accumulated dropped packets for each type of traffic in the two sinks belonging to different domains.

As seen in Table 11, once the policies are applied the system stops dropping EF packets, reduces the AF41 packets drop to 4.79 % and the AF21 packets drop to 6.03 % and increases to 33.68 % (average of both sinks) the BE packets drop. These results are coherent to the policy applied since the EF queue has absolute priority and AF41 has a lower discarding probability than AF21.

	Sink-1		Sink-2	
	Before policies are applied	**After policies are applied**	**Before policies are applied**	**After policies are applied**
EF	14.97 %	0 %	16.46 %	0 %
AF41	22.88 %	4.79 %	-	-
AF21	-	-	22.43 %	6.03 %
BE	13.67 %	33.86%	17.5 %	33.50%

Table 11: Percentage of dropped packets (inter-domain)

These results are correct according to the policies implemented in the edge and core routers. Traffic was compliant to the admission traffic policies in the edge routers and the percentage of droping packets in the core routers conform the policies. The policy provided maximum priority for EF packets and provided to AF41 traffic a smaller drop probability compared to AF21.

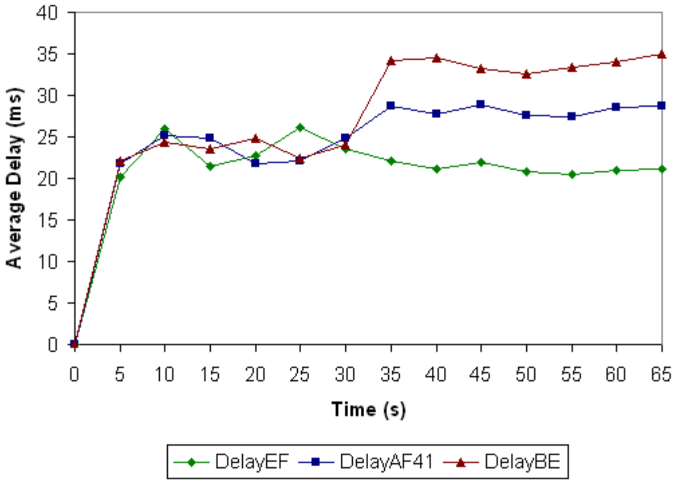


Figure 45: Average packet delay destined to sink-1 before and after the policies are applied

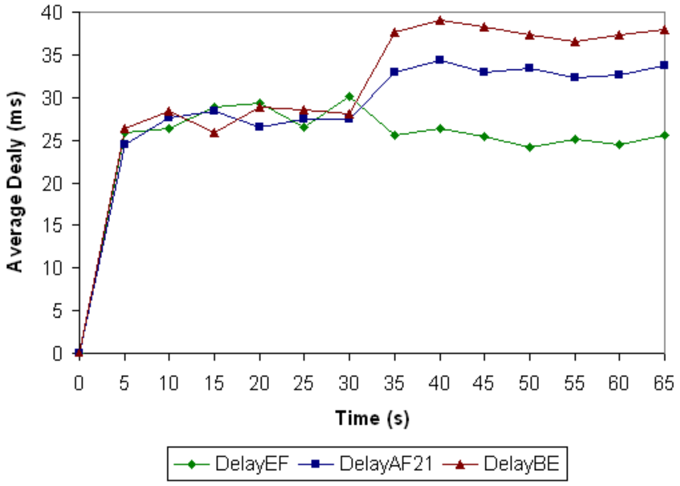


Figure 46: Average packet delay destined to sink-2 before and after the policies are applied

Figure 45 and 46 illustrate the behavior of the packet delay parameter in 5-second intervals averages within the same test. Both graphs show similar behaviors between them even though sink-2 one domain further and therefore presents more delay. These results are also very similar to those shown in Figure 39 which is logical because traffic conditions are very similar and QoS policies are the same. When policies are applied it can be observed how traffic AF21 is a little bit more delayed than AF41 and therefore it approximates a little bit more to BE delaly results in sink-2. These results are coherent with the policies applied the same way as in the intra-domian results. This is beacuse EF packets have always maximum priority. Once EF packets are shaped at the first edge router they observe virtually no queuing delay. On the other hand AF and BE packets are queued with statistical priorities and drops. Furthermore, BE packets have minimum priority.

Thefore the results validated the overall functioning of the whole BBv6 architecture. Moreover, there existed several problems in the testing because of the flaws in the timing synchronization between the nodes with the NTU protocol and the poor response of the traffic generation softwares available (Iperf and DITG) which made the tests take much longer than necessary. Besides, the policy management, the performance of the policy-delivery system, the COPS-PR protocol, the GNU/Linux testbed and the modules implemented in each node, which had already been validated in the previous section, these results affirm how well the system behaves in a multi-domain environment and validate the SIBBSv6 protocol. These results are therefore presented as a contribution of this thesis [203].

3.8. Chapter summary

The bandwidth brokers provide centralized management of the DiffServ QoS in multi-domain environments allowing seamless end-to-end QoS for end-to-end multimedia communications. A bandwidth broker for native IPv6 networks, named BBv6, has been introduced, described and evaluated as stated in the CSI-RHET project.

The state of the art conducted have shown that, even though there are many scientific papers proposing bandwith brokers through simulations, there are very few bandwith brokers proposals really implemented. This state of art has been the basis of the BBv6 development which has incorporated the COPS and COPS-PR standard protocols, as well as the new SIBBv6 protocol for the policies delivery.

This new BBv6 device provides the operators and the researchers with a bandwidth broker to centrally manage native IPv6 DiffServ domains which permits the delivery of network policies through the COPS-PR protocol, a standardized intra-domain communication protocol that optimized policy delivery, to IPv4 and/or IPv6 domain nodes. The implementation of the inter-domain signalling among peer bandwidth brokers in different IPv6 DiffServ domains has been done with the SIBBSv6 protocol, a new version of the SIBBS protocol for IPv6 which solves some of its flaws demosntrated by literature. With SIBBv6 resource reservation is now carried out through SLAs/SLSs and the finalisation of the acceptance notification of all the remote BBv6 along the path is now a requirement before sending the data in the signalling process.

The BBv6 has been programmed with JAVA for multiplatform support and it is open source. Its modular design allows specialized modules for each of the tasks it is in charge of and enables adaptability for new technologies and protocols in the future.

The system has been successfully evaluated with a three IPv6 DiffServ domain testbed, with three BBv6 managing each of these domains, which has been designed and implemented to evaluate the operation and performance of the architecture. All the nodes of the DiffServ domain, implemented with GNU/Linux, support the Traffic Conditioning mechanisms and the PHB stated in the DiffServ architecture.

Two blocks of tests have been carried out. The first block has evaluated the intradomain policy delivery protocol performance and the second block has evaluated the interdomain protocol performance. The results show how the BBv6 has successfully accepted the remote definition of policies, their storage, their interaction with other peer BBv6, their translation to COPS-PR protocol's PIB system, their delivery to the DiffServ domain nodes with COPS-PR and finally their correct mapping so that the GNU/Linux O.S. can successfully prioritize all the premium traffic and most of the AF traffic although the latter cannot be guaranteed.

These contributions to the thesis have been presented in [202] and [203].

4. QoS Control in ITU-T's NGN Architecture

4.1. Introduction

The International Union for Telecommunications, Telecommunications Sector, (ITU-T) has developed a generic end-to-end architecture for the network architecture and control procedures of Next Generation Networks (NGNs) which comprises core and access networks [116]. This architecture incorporates the IP Multimedia Subsystem (IMS) [117] to support session-based services, and other services based on Session Initiation protocol (SIP). The ultimate objective of this new architecture is to provide a unique service infrastructure for the management of new services and multimedia communications within diverse NGNs.

This architecture provides several benefits to the communications industry. There will be cost reductions as network infrastructure and systems will be shared and there will be a simplification of the operations and management, thus lowering the operational expenditure (OPEX) given that operation platforms, maintenance and training will be integrated. Finally, it facilitates the creation of new services and therefore revenue will increase. On the other hand, the operators are required to support flexibility to incorporate new services and they have to react quickly to new ones, to support survivability to allow service assurance in case of failure, QoS to guarantee clients SLAs for different traffic mixes, conditions and overload, and finally to support interoperability across networks to allow end-to-end services for flows in different network domains.

The establishment of a mechanism for end-to-end QoS control management in a heterogeneous NGN infrastructure is a complex task. The coexistence of multiple QoS technologies and different provider domains must be taken into consideration. One of the key aspects will be the study of the interconnection of the different providers and technologies. Therefore one of the main implementation difficulties will be in the efficient and dynamic exchange of policies between the different domains, as defined in their respective SLAs. In addition, this exchange must be secure when trust, which refers to authentication, confidentiality and integrity, is not assured between domains of different providers.

Various organizations have developed specifications related to NGNs. The IEEE has focused on developing solutions for layer 2 issues, while the IETF has concentrated on layer 3 issues. ITU-T and ETSI are developing network architecture and control procedures [187]. Different agents have also taken part in the creation of the resource and QoS control architecture for NGNs: CableLab for HFC cable access networks, the DSL forum for DSL access networks, the 3rd Generation Partnership Project (3GPP) for mobile access networks and ETSI-TISPAN for generic access networks which work independently from transport technology while ITU-T has developed a generic end-to-end architecture which summarizes the aforementioned. In order to offer a differentiating factor, every organization re-using other organizations' architectures, have renamed the interfaces and functional blocks, even when the architecture carries out exactly the same functions. This has resulted in a chaos of nomenclature (Figure 47).

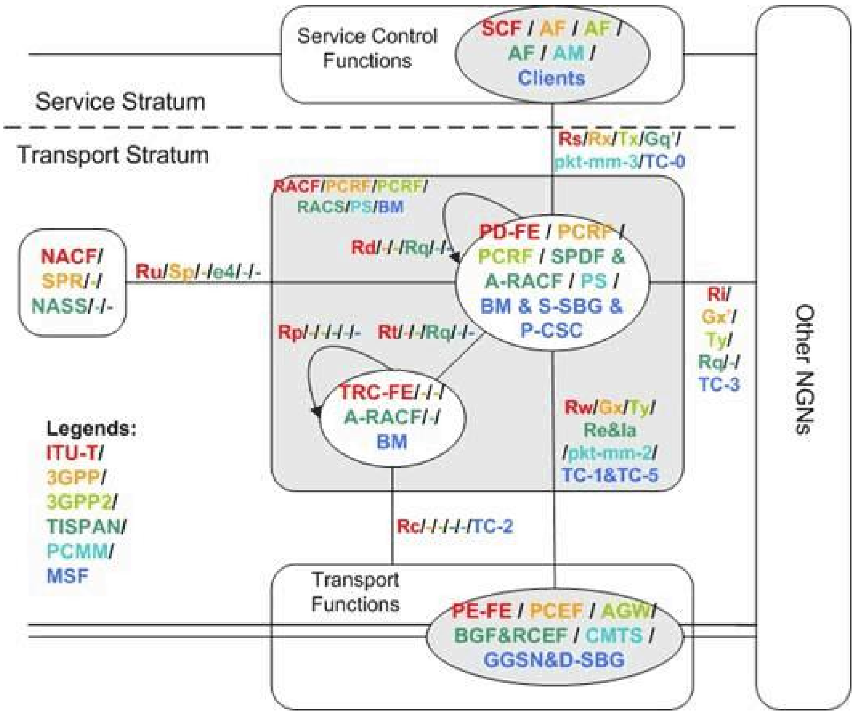

Service Control Functions
SCF / AF / AF / AF / AM / Clients
Service Stratum
Transport Stratum
Rs/Rx/Tx/Gq'/ pkt-mm-3/TC-0
RACF/PCRF/PCRF/ RACS/PS/BM
Rd/-/-/Rq/-/-
PD-FE / PCRF / PCRF / SPDF & A-RACF / PS / BM & S-SBG & P-CSC
NACF/ SPR/-/ NASS/-/-
Ru/Sp/-/e4/-/-
Rp/-/-/-/-/-
Rt/-/-/Rq/-/-
TRC-FE/-/-/ A-RACF/-/ BM
Ri/ Gx'/ Ty/ Rq/-/ TC-3
Other NGNs
Rw/Gx/Ty/ Re&Ia /pkt-mm-2/ TC-1&TC-5
Legends:
ITU-T/
3GPP/
3GPP2/
TISPAN/
PCMM/
MSF
Rc/-/-/-/-/TC-2
Transport Functions
PE-FE / PCEF / AGW/ BGF&RCEF / CMTS / GGSN&D-SBG

Figure 47: Nomenclatures of resource and QoS control architectures for NGNs

This chapter sheds some light over the confusing nomenclature, and enhances the BBv6 device into a QoS broker compliant with the ITU-T requirements.

4.1.1. Objectives

The main objective of this chapter is to enhance the BBv6 architecture presented in the previous chapter with the aim of meeting the ITU-T's NGN requirements and, thus, developing an end-to-end QoS control architecture able to provide QoS with end-to-end multimedia communications. In order to achieve this, the first objective is to study the state of the art of the ITU-T NGN architecture development status and understand the resource and admission control functions of the QoS control framework.

The second objective is to develop a QoS Broker (QoSB) device enhancing the BBv6 device according to the ITU-T NGN requirements. One of the key factors will be the determination of the communication protocol for the inter-domain relationship between peer QoS brokers. The BBv6 was built in modular basis and this will ease the implementation of the new protocols because only the communications modules might be affected.

The final objective of this chapter is to validate the QoS Broker architecture. The performance evaluation will use the testbed developed in the previous chapter for the CSI-RHET project. The success of this third phase of the thesis will be measured by the correct functioning of the QoS Broker architecture, built according to the aforementioned requirements.

4.2. ITU-T Next Generation Networks architecture

According to ITU-T "a Next Generation Network (NGN) is a packet-based network able to provide services including Telecommunication Services and able to make use of multiple broadband, Quality of Service-enabled transport technologies and in which service-related functions are independent from underlying transport-related technologies. It offers unrestricted access by users to different service providers. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users" [115].

ITU-T's NGN-Global Standards Initiative (NGN-GSI) was first created in 2003, but was in 2004 when it created a Focus Group on NGN (FGNGN) to address the urgent need for an initial suite of global standards for NGN. In November 2005, the results of the FGNGN (Release 1) provided the building blocks on which the world's systems vendors and service providers could start to make the shift to NGN. The ITU-T's NGN-GSI encompasses all NGN work across ITU-T Study Groups. These ITU-T activities are related to the establishment of architectures, interface specifications, and implementation guidelines in the form of ITU-T standards (Recommendations) for the realisation of NGN.

The purpose of the ITU-T's NGN-GSI is to provide a visible single location for information on and the coordination of the development of NGN standards. These are the detailed standards necessary for NGN deployment to give service providers the means to offer the wide range of services expected in NGN.

It has been decided to use the ITU-T architecture as a base for enhancing the architecture presented in the previous section because it proposes a global architecture for managing QoS in an end-to-end basis and comprises core and access networks. This new NGN architecture presents a stratified division between the service plane in the upper layer, including the session control plane in the intermediate control sub layer, and the transport plane in the lower layer (Figure 48). In this latter case, the resource and admission control functions have been defined in the RACF entity [119], which is capable of managing the end-to-end QoS through both core and access heterogeneous networks. This architecture is presently under definition and therefore some of the reference points are still to be specified.

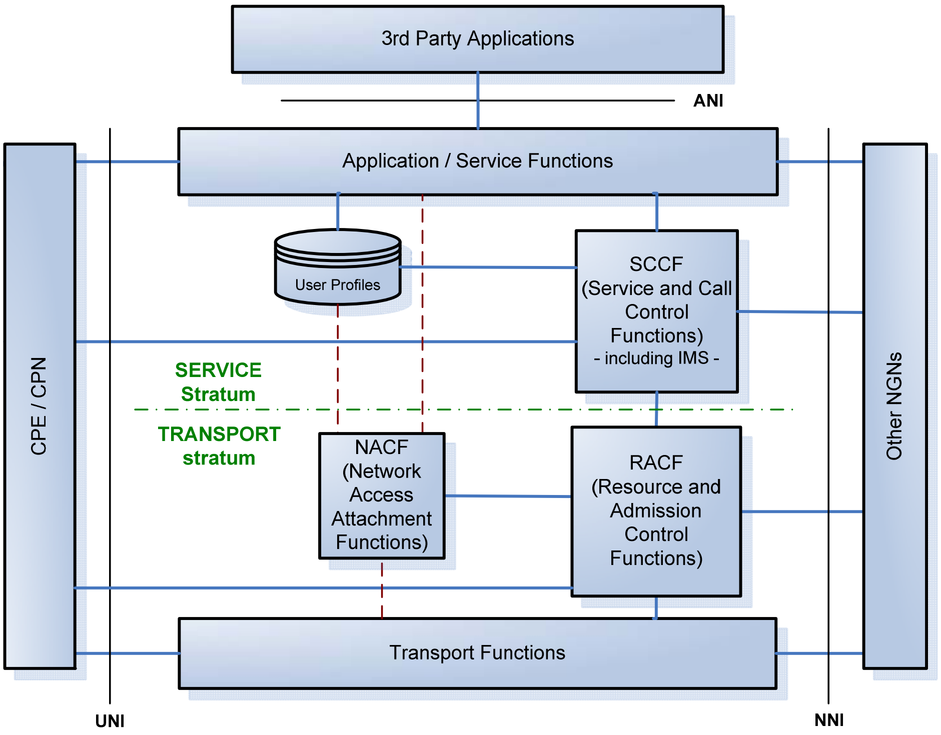


Figure 48: ITU-T NGN framework architecture

The NGNs are based on IETF Internet technologies including IPv6 and Multiprotocol Label Switching (MPLS). At the application level they are based on Session Initiation Protocol (SIP). Softswitch devices are used for voice applications to control ToIP calls and to create an interface to the existing telephone network, PSTN, through Signalling Gateways and Media Gateways. But the most important technology is the IP Multimedia Subsystem (IMS).

The ITU-T has elected the IMS (IP Multimedia Subsystem) architecture, developed by 3GPP, as a support of session-based services and other session initiation protocols (SIP) [117]. This architecture uses the reference points defined in the transport subjacent infrastructure to manage the negotiated QoS through the session signalling and the flow control at a network level.The QoS management architecture proposed by ITU-T for NGNs is completely integrated and interoperable with IMS.

4.2.1. IP Multimedia Subsystem

The IP Multimedia Subsystem (IMS) architecture is the base of NGN development. This architecture has been defined by the 3GPP project [230] which is in charge of the supervision of the standard development related to 3G technology. IMS is not intended to standardize applications but rather to aid the access of multimedia and voice applications from wireless and wireline terminals. IMS was initially incorporated into the mobile network core but has since been expanded to all IP networks.

The IMS is a collection of core network functional entities for the support of SIP-based services [117]. IMS supports the registration of the user and the terminal device at a particular location in the network. As part of registration, IMS supports authentication and other security arrangements. The services supported by IMS may include multimedia session services and some non-session services such as Presence services or message exchange services.

As illustrated in Figure 49, IMS provides a horizontal control layer that isolates the transport layer (e.g. routers) from the service layer (applications). In the IMS layer there are different SIP servers and proxies which have been aggregated in an entity named CSCF (Call Session Control Function). This key functional entity executes three different roles. The P-CSCF (proxy CSCF) is the entry point for all the SIP sessions in the system coming from the end user terminals and provides protection, security and resources control for the transport subsystem. The S-CSCF (Serving CSCF) registers and authenticates users, and routes the user sessions. Every registered user is assigned an S-CSCF. The I-CSCF (Interrogating CSCF) is a supporting entity which helps route signalling.

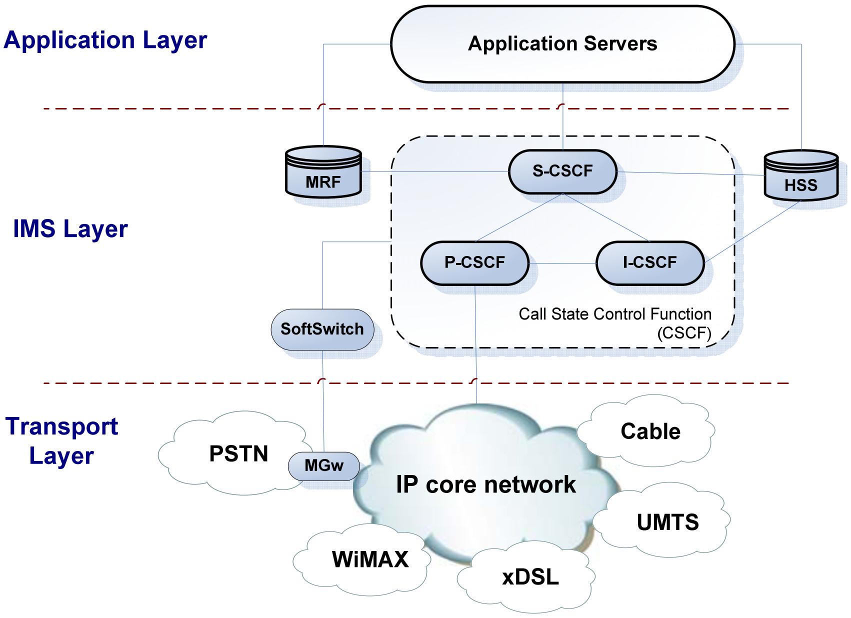


Figure 49: IMS architecture of the 3GPP

To ease the integration with the Internet, IMS uses IETF protocols whenever possible, as part of the development agreement between the 3GPP and the IETF. The main IP protocols in IMS are SIP and IPv6. The session control is done by the IMS call control protocol based on SIP and SDP. The IMS has been defined from the beginning as a system completely based on IPv6. Besides these two basic protocols, IMS uses other IETF protocols: RTP (Real Time Protocol), RTCP (Real Time Control Protocol), RSVP (Resource Reservation Protocol), DiffServ, COPS (Common Open Policy Service), Diameter or Megaco.

In December 2009, the SIP IPv6 sub-working group at the v6PC released a trial version of an IMS IPv6 conformance test [106].

4.2.2. RACF architecture

In order to provide efficient end-to-end QoS control management, the NGN-GSI architecture has adopted the policy-based network management proposed by the IETF Policy Framework Working Group [98]. The aim is to provide a centralized management system to control the end-to-end QoS of the multimedia sessions. The ITU-T has consequently developed the QoS management sub layer and has defined the RACF entity providing the SCF with an abstract vision of the network infrastructure (Figure 50) [119]. RACF carries out the policy-based physical resource control, establishes the availability of these resources, decides on admission and applies the controls required to enforce the accomplishment of the policy decisions. Even though different international organizations (ITU-T, ETSI-TISPAN, MSF, etc.) have proposed many alternative protocols to the network policies delivery in use in the Rw, Rc and Rn interfaces, industry has not opted for any of them.

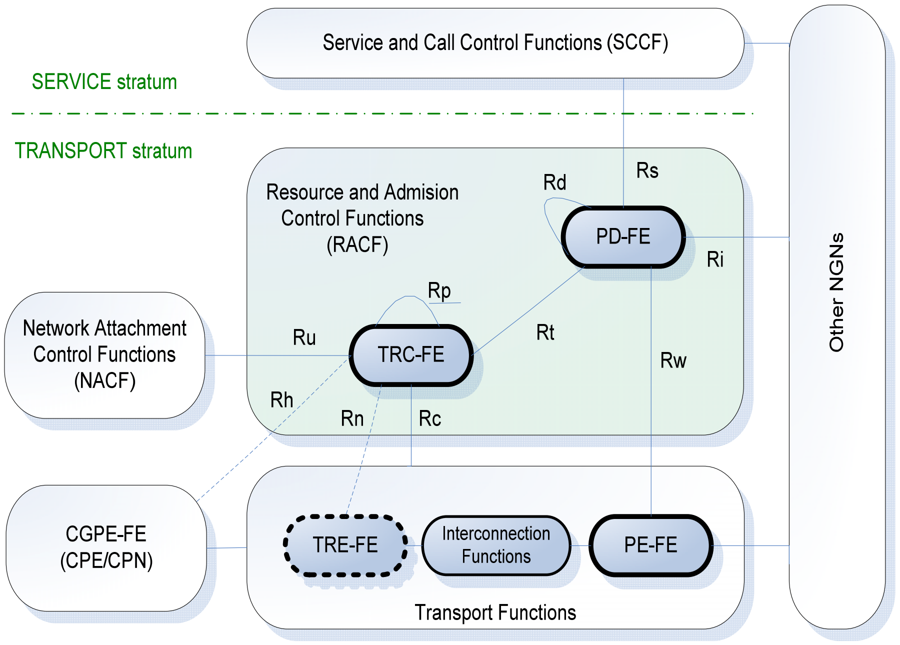

Service and Call Control Functions (SCCF)
SERVICE stratum
TRANSPORT stratum
Resource and Admision Control Functions (RACF)
Rd
Rs
PD-FE
Ri
Rp
Ru
TRC-FE
Rt
Rw
Network Attachment Control Functions (NACF)
Rh
Rn
Rc
CGPE-FE (CPE/CPN)
TRE-FE
Interconnection Functions
PE-FE
Transport Functions
Other NGNs

Figure 50: RACF architecture proposed by the ITU-T

The Policy Decision Functional Entity (PD-FE) takes the final decision over the resource and admission control and delivers it to the corresponding Policy Enforcement Functional Entity (PE-FE) through the Rw interface. When this study took place in autumn 2007, the ITU-T was studying three alternatives for this interface, specified in the Q.3323.x sub-series [112]. Two of them had already been approved COPS-PR and H.248. The third, DIAMETER, was awaiting approval. In April 2010, the three of them are approved (Table 12).

The Transport Resource Control Functional Entities (TRC-FE) deals with the control of the resources which depend on transport technology. These entities are responsible for preserving and maintaining the network topology and resource database (NTRD) of each subdomain. The Rc interface is used to check the network topology and the status of network resources. The TRC-FE assigns resources to each QoS requesting flow. In 2007 the ITU-T had approved two alternatives for this interface [119], which had been specified in the Q.3324.x sub-series [113]: COPS-PR and SNMP. The specific details of this interface were still under study. In April 2010, both of them are approved (Table 12).

According to [119] the scope and functions of the Rn interface are also still under study in 2010, though it does clarify that one of the functions of the TRC-FE entity is to assign the network resources for its application. Given that the functionalities of the Rc interface have been already defined, although not in full detail, the only interface able to accomplish the aforementioned functions is the Rn interface. It seems logical to think that new protocols will be proposed to address network policy delivery to meet requirements. However, ITU-T has not made any formal proposals. Some of the protocols which meet the requirements and have been used by ITU-T in other interfaces include SNMP and COPS-PR.

SCFs, RACFs and transport functions may collaborate with other NGN-GSI architectures in order to provide services to nearby domains and therefore achieve end-to-end QoS. When the collaboration takes place at the RACF level it is done through the Ri interface. This interface was also under development in 2007 and only some basic requirements had been released. In 2010 the official status is the same but staff of the SG.11 [122] has unofficially informed to the author of this thesis that the Q.3307.1 recommendation on the Ri interface is pre-published, it will be officially released soon and the protocol for the Ri interface will be a QoS variation of DIAMETER still under development. This fact will be later on detailed in this chapter.

Even though the transport resource control in NGNs includes QoS control, NAPT and firewall control, as well as NAT traversal, the scope of this thesis is limited to QoS control.

Interface	Supporting Entities	Protocol (in 2007)	Protocol (in 2010)	Rec. No.
Rs	SCE, PD-FE	DIAMETER	DIAMETER	Q.3321.1
Rp	TRC-FE, TRC-FE	RCIP	RCIP	Q.3322.1
Rw	PD-FE, PE-FE	COPS-PR	COPS-PR	Q.3323.1
		H.248	H.248	Q.3323.2
		DIAMETER (pre-published)	DIAMETER	Q.3323.3
Rc	TRC-FE, T-FE	COPS-PR	COPS-PR	Q.3324.1
		SNMP	SNMP	Q.3324.2
Rt	PD-FE, TRC-FE	-	DIAMETER	Q.3325.1
Ru	PD-FE, NACF	-	DIAMETER	Q.nacp.Ru
Rd	PD-FE, PD-FE	-	DIAMETER (pre-published)	Q.3306.1
Ri	PD-FE, PD-FE	-	DIAMETER (pre-published)	Q.3307.1
Rn	TRC-FE, TRE-FE	-	Interface for further study	-
Rh	TRC-FE, CGPE-FE	-	- (pre-published)	Q.3308.1

Table 12: RACF protocol recommendations

4.3. Intra-domain control of QoS Resources

The RACF entity defines two possible scenarios for the QoS control which are based on the type of sessions established by the user. Depending on the QoS signalling capacity of the terminals (CPE) that initiate the sessions and the access technology, the QoS resources control may be in push mode or in pull mode (Figure 51).

The push mode is initiated through the session signalling (with the SIP protocol) when QoS resource control signalling does not exist in the CPE. The SCCF is responsible for deriving the QoS needs of the requested service and sending the request to the RACF for QoS authorization and reservation and then finally sending the QoS policy to the network transport equipment (PE-FE). In this mode every time the PD-FE (in the RACF) receives a QoS resource request from a SCF, it decides on authorization control and resources and then autonomously orders the transport functions to carry out these decisions, sending the QoS policy to the network transport equipment (PE-FE). This mode is employed by terminals without QoS negotiating capacity (type-1 CPE) or by those with only service stratum negotiating capacity (type-2 CPE).

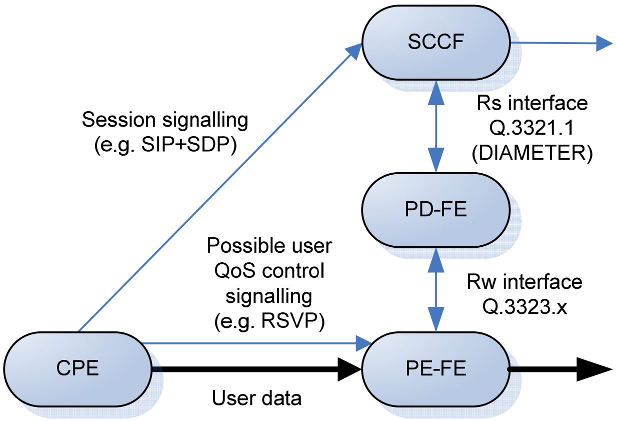


Figure 51: Push or pull operation depending on the type of user terminal

The pull mode is initiated when user QoS control signalling does exist in the CPE and the transport functions require signalling to perform a flow (e.g. RSVP or NSIS). In this mode it is the PE-FE which sends a QoS resource request to the PD-FE through the Rw interface, so that the RACF may take the appropriate authorization decision and replay with the final policy decision to be applied. This mode is used by terminals that explicitly request the QoS resource request through path-coupled QoS signalling (type-3 CPE).

In autumn 2007, the two protocols proposed by the ITU-T to manage the Rw interface in core networks, DIAMETER [19] and COPS-PR, had similar problems due to the natural client-server role (pull mode), which was inverted when the push mode was in use. In the original definition of the COPS-PR protocol, it works efficiently in pull mode but not so well in push mode. On the other hand, the original definition of the DIAMETER protocol also works correctly in pull mode but in this case the push mode is not contemplated. Two draft specifications have been proposed to solve this problem within both protocols [219][229]. In these drafts two extensions have been proposed to solve the problem and legitimate their utilization in push mode, making them as efficient as in the pull mode in both cases. In 2010, as aforementioned, a new variation of the DIAMETER protocol for managing QoS is being developed.

Therefore, while most of the protocols defined in RACF interfaces and routing elements of the transport layer are still under definition, the common protocol for all of them will probably be COPS-PR. An additional factor to consider is that the COPS-PR which manages policy delivery on a native level is already supported by most IP-IP gateways (routers) in present-day networks. When the QoS control in the NGN-GSI architecture is applied into the network devices to manage the domain's QoS, these devices have to support the COPS-PR protocol as they will perform the transport functions and will therefore be the PE-FE and the TRC-FE.

4.3.1. COPS-PR use in the ITU-T's NGN-GSI architecture

The COPS-PR protocol has been chosen because most of the protocols defined in RACF interfaces and routing elements of the transport layer are still under definition and it will probably become the common protocol for all of them. An additional factor to consider is that COPS-PR, which manages policy delivery on a native level, is already supported by most IP-IP gateways (routers) in present-day networks.

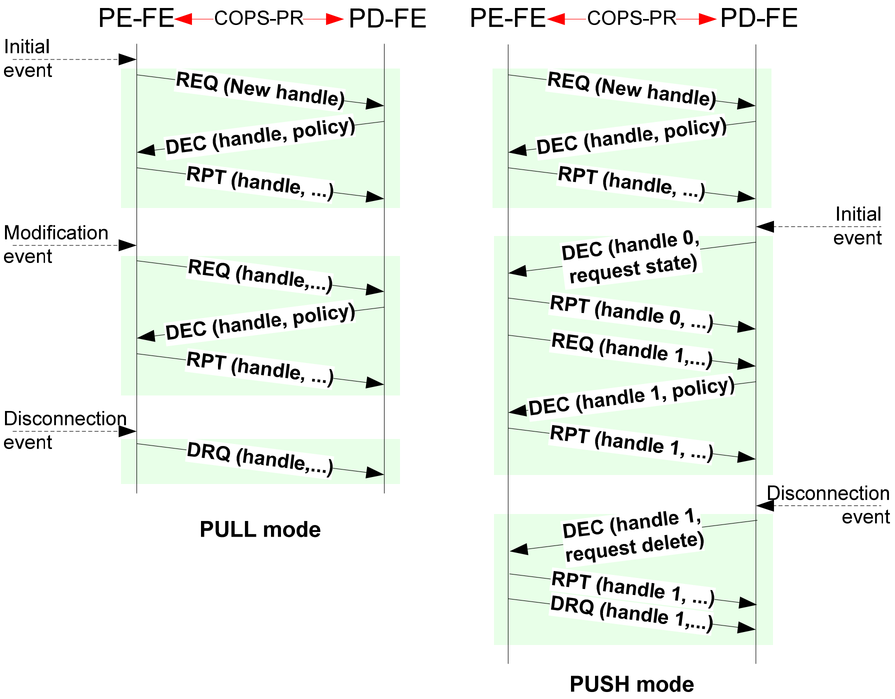


Figure 52: COPS-PR signalling in pull mode and push mode

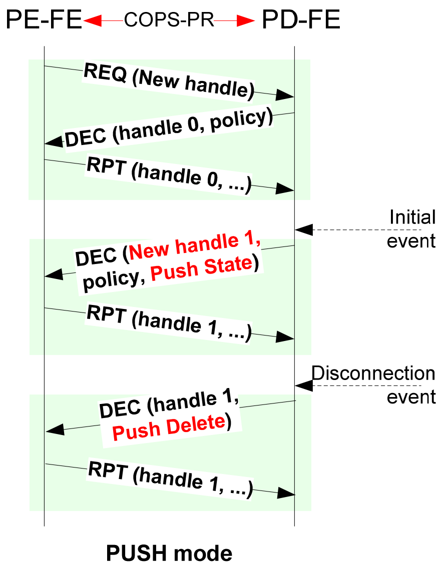


Figure 53: Proposal of COPS-PR modification in push mode

The COPS protocol was originally design to work in pull mode. Therefore the message exchanges in this mode are efficient and support a close relationship between the operations of the mentioned protocol and the state associated to the particular events detected by the PE-FE (NGN terminology for the PEP). Within the COPS-PR protocol, the provisioning model for the DiffServ model, the PD-FE proactively sends the network policies to be applied to the PE-FE. As stated before, when COPS-PR is applied to the RACF entity, the protocol works effectively in pull mode, but not so well in push mode, even though it does work correctly. In Figure 52 it can be seen how this latter mode incorporates appended messages between the PD-FE (NGN terminology for the PDP) and the PE-FE when the states of the different events which appear (detected by the PD-FE) are associated to the specific operations of the protocol.

The proposal of modification of the COPS-PR protocol in [219] to optimize the push mode operation consists in removing some of the messages and creating a new value in the "New Decision" flag (Figure 53). The problem with this proposal is that it involves modifying the protocol, and therefore loses interoperability, which is one of the main advantages of COPS-PR.

It has been decided to use the standard COPS-PR protocol in the QoS broker because of the aforementioned reasons and because it is already implemented in the BBv6.

4.4. Inter-domain control of QoS resources

ITU-T defines the functions and processes needed to support QoS between different provider interfaces. In the inter-domain signalling between providers there are three ways to manage end-to-end QoS control: Through the application layer, through the Ri interface (RACF) and through path-coupled QoS signalling (e.g., RSVP-like). This is why under certain scenarios the intervention of the Ri interface may not be needed.

The use of the Ri interface has been proposed in the event of there being intermediate providers where the SCF cannot operate with the correspondent PD-FE (or if it has no application functions), where path-coupled QoS signalling does not exist or when the SCF of a core network can only communicate with the PD-FE of the access networks through its own core network PD-FE.

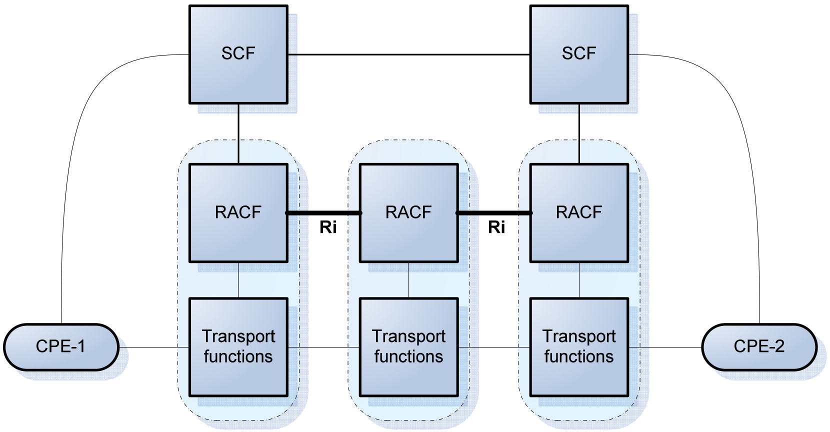


Figure 54: RACF communications between operators

4.4.1. Dynamic SLS negotiation

Simple and automatic network configuration is a starting point for dynamic SLAs. This can decrease the time needed for SLA negotiation and deployment. These automated configuration methods improve the network providers' flexibility in support of the customers' dynamic needs, while also reducing the excess capacity on the networks. Dynamic SLAs may match the service levels with the needs of the client as well as the capability of the network. The provisioning of end-to-end QoS assumes the negotiation of mutually acceptable SLAs. The SLA specifies the

QoS level required by a given service and the operations to apply within the network to handle the corresponding traffic flows. An SLA can be negotiated between a user and its network, as well as between various networks.

Most of the service level negotiation process has been usually carried out manually. Such an SLA is considered static. Frequent modification is almost impossible, due, for example, to the potential interactions between SLAs and the constraints at the enforcement points. When modification is needed, the customer may require the assistance of the network provider. The provider, in turn, may need to reengineer part of a system to ensure the mandated service levels. A dynamic SLA, on the other hand, could be modified more easily. Since there is no formal definition of dynamic SLA to date, it is commonly defined the dynamic SLA as a contract between client and network provider in which a subset of parameters can be changed over time. In order for the service level negotiation to be dynamic on a large scale, a protocol for SLS negotiation is mandatory.

NGNs will require a high flexibility in resource allocation. This demand implies dynamic SLS negotiation in which service levels are negotiated dynamically and resources are allocated on demand.

Dynamic service negotiation protocols

In 2007 the ITU-T had not proposed any Ri interface protocols, although from the defined functionalities it was expected that those to be proposed would be dynamic service negotiation protocols. A comparative study of the dynamic network policies negotiation protocols proposed by different authors to be used in NGNs is presented by Sarnagan et.al. in [181]. These protocols include: RNAP [214], SrNP [192], DSNP [25], QoS-NSLP [132], QoS-GSLP [3] and COPS-SLS [153]. All of them negotiate SLS (Service Level Specifications) and even though there many types exist, these protocols negotiate mainly QoS SLS. The conclusion is that none of them are clearly better than the rest because each one has advantages and disadvantages (Table 13).

Other protocols have been proposed for inter-domain communication but they do not accomplish ITU-T requeriments as none of them negotiate SLS. Some examples are RSVP [14], BGRP [134][162], DARIS [13], COPS-DRA [180] or SIBBS.

Given that the proposed protocol to be implemented in the intra-domain interfaces for the QoS control is COPS-PR, and therefore is already implemented in the PD-FE, it was proposed for this study in 2007 the use of COPS-SLS as the protocol for the Ri interface. This election results in an internal resource optimization of the PD-FE and facilitates the interoperability between the resource assignation protocols of different interfaces.

However, a scalability problem appears in the SLS inter-domain negotiation when there are an important number of domains to interconnect. In [133] an architecture to solve the problem is proposed by applying a hierarchical structure to the PD-FE and creating local PD-FEs. The application of COPS-SLS in its hierarchical mode would permit the implementation of such a solution. It must be mentioned that this factor is not contemplated in the study by [181].

The protocol implemented in the BBv6 was SIBBSv6, and even though it has been proposed over COPS-SLS in [148] when scalability is not a factor it does not fit the ITU-T requeriments.

Feature	Protocol					
	RNAP [214]	SrNP [192]	COPS-SLS [153]	DSNP [25]	QoS-NSLP [132]	QoS-GSLP [3]
Primary negotiation	Yes	Yes	Yes	Yes	Yes	Yes
Generic QoS architecture	Yes	Yes	Yes	Yes	Yes	Yes
Link layer transparency	Yes	Yes	Yes	Yes	Yes	Yes
Extends existing protocol	Yes	Yes	Yes	No	Yes	Yes
Lightweight	No	No	No	Yes	No	Yes
Reduced signalling	No	Maybe	Yes	Yes	No	Maybe
SLS format transparency	No	Yes	No	No	Maybe	No

Table 13: Comparative study of protocols for dynamic service negotiation [181]

4.4.2. COPS-SLS protocol

In this thesis it has been proposed the use of COPS-SLS (COPS for Service Level Negotiation) [153][154] as the protocol for the Ri interface. The COPS-SLS protocol is not a standardized extension of COPS which exploits the characteristics of COPS in service level negotiation and it is able to dynamically negotiate the service levels between the users and their domain and between domains. The three basic messages of COPS protocol (request, decision, and report) can be used to support all required basic negotiation operations and cope with the requirement of a reliable negotiation. In fact, COPS-SLS supports the same basic procedures of COPS for connection stablishment, connection termination and fault tolerance.

COPS-SLS comprises two phases, configuration and negotiation, as illustrated in Figure 55. In the configuration phase, the SLS PD-FE uses the Provisioning model to tell the client how to request a level of service. It supplies the client with information about the negotiation mode and/or the time interval to renegotiate. The SLS PE-FE installs this configuration and sends a report to the SLS PD-FE. After successfully installing the configuration supplied by the SLS PD-FE, the SLS PE-FE can start the negotiation phase using the Outsourcing model. The SLS PE-FE sends a request for its desired level of service. The SLS PD-FE can accept the request, reject the request, or propose another level of service to the client. At any time, the SLS PD-FE can send an unsolicited decision to change the parameter values of the configuration phase or to degrade a service level in case of heavy congestion in the domain. Each REQ message is identified by the client-handle value. The client-handle value in the DEC message identifies the request to which it replies. In the RPT message, the client-handle value identifies the decision on which it reports.

The organization of COPS-SLS in two phases (configuration and negotiation) makes the negotiation dynamic and configurable. In COPS-SLS, the service templates offered by the provider can be provisioned in the configuration phase using predefined SLSs. The negotiation phase is paced by the configuration phase. When the network detects a problem that influences current SLA dynamism, the PDP can slow the rhythm of the negotiation down by sending an unsolicited decision with the *slsNegoMaxInt* setting to a larger timescale the ISP finds convenient. In the configuration phase, the attribute *slsNegoMaxInt*, in the JAVA class *slsNgoEntry*, indicates the maximum time interval to renegotiate. This attribute reflects the dynamicness of the SLA because it specifies the timescale in which the client can modify an already negotiated SLS. Note that the decision in the configuration phase does not influence the level of service of a client but changes the provider's manner of negotiation with this client.

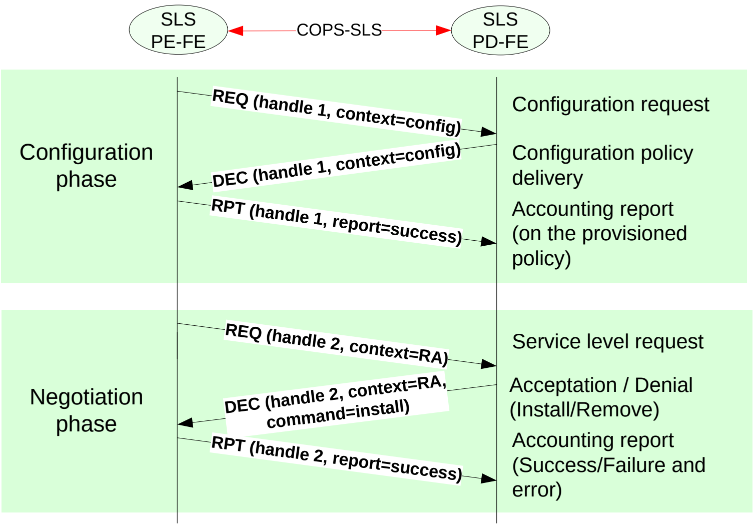

SLS
PE-FE
COPS-SLS
SLS
PD-FE
Configuration phase
REQ (handle 1, context=config)
DEC (handle 1, context=config)
RPT (handle 1, report=success)
Configuration request
Configuration policy delivery
Accounting report (on the provisioned policy)
Negotiation phase
REQ (handle 2, context=RA)
DEC (handle 2, context=RA, command=install)
RPT (handle 2, report=success)
Service level request
Acceptation / Denial (Install/Remove)
Accounting report (Success/Failure and error)

Figure 55: COPS-SLS signalling sequence

COPS-SLS provides SLS management for inter-domain communication in a multi-domain environment. Therefore, as it is a client-server architecture, the QoS broker will have the double role of PD-FE and PE-FE. Figure 56 presents a message sequence chart of the negotiation between two QoS brokers managing two domains. In this example, domain A is assumed to be a corporate network providing QoS and domain B is a national ISP.

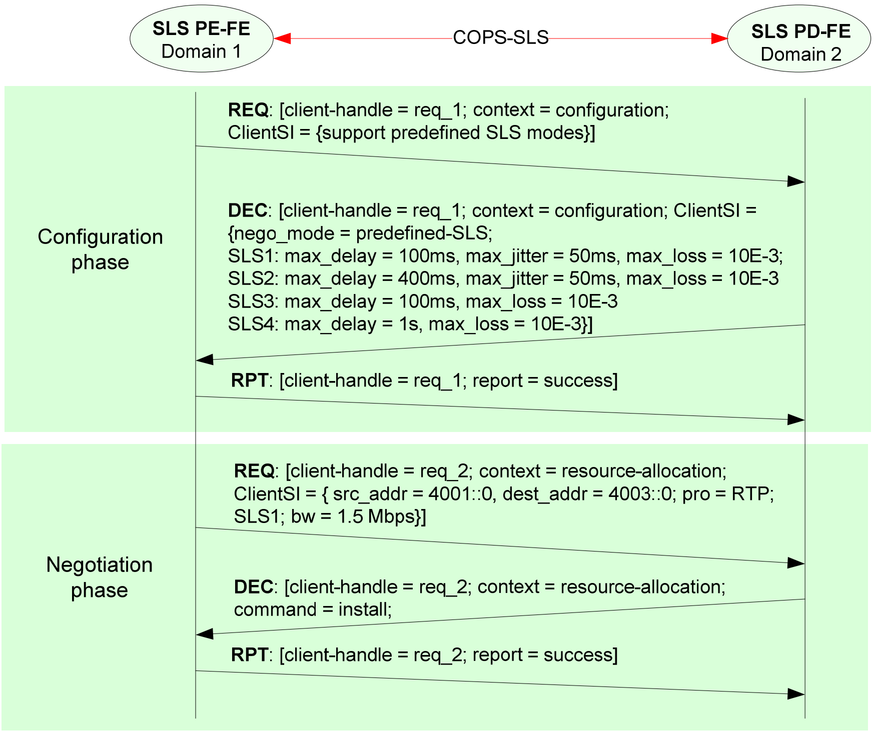

SLS PE-FE
Domain 1
COPS-SLS
SLS PD-FE
Domain 2
Configuration phase
REQ: [client-handle = req_1; context = configuration; ClientSI = {support predefined SLS modes}]
DEC: [client-handle = req_1; context = configuration; ClientSI = {nego_mode = predefined-SLS;
SLS1: max_delay = 100ms, max_jitter = 50ms, max_loss = 10E-3;
SLS2: max_delay = 400ms, max_jitter = 50ms, max_loss = 10E-3
SLS3: max_delay = 100ms, max_loss = 10E-3
SLS4: max_delay = 1s, max_loss = 10E-3}]
RPT: [client-handle = req_1; report = success]
Negotiation phase
REQ: [client-handle = req_2; context = resource-allocation; ClientSI = { src_addr = 4001::0, dest_addr = 4003::0; pro = RTP; SLS1; bw = 1.5 Mbps}]
DEC: [client-handle = req_2; context = resource-allocation; command = install;
RPT: [client-handle = req_2; report = success]

Figure 56: COPS-SLS interdomain negotiation with predefined SLS

In the configuration phase illustrated in Figure 56, domain B sends a decision message to provision domain A four predefined SLSs called SLS1, SLS2, SLS3, and SLS4 with different constraints concerning QoS parameter values. The PEP can request a service level conforming to one of the predefined SLSs provided by the PDP. In the Negotiation phase, domain 1 requests a service level for an aggregate sharing 1.5 Mbps between all Real Time Transport (RTP) flows from domain 1 towards domain 3 with max_delay = 100 ms, max_jitter = 50ms, and max_loss = 10E-3. This aggregate is identified by the subnet addresses of the source domain (4001::0), the subnet addresses of the destination domain (4003::0), and the protocol value (RTP). The service ID is supported, and therefore the REQ message simply indicates 'serviceID = SLS1' in place of specifying delay, jitter and loss ratio values.

The COPS-SLS architecture is designed as a flexible buildine block to achieve end-to-end service level negotiation and QoS management and therefore allows 3 basic negotiation models (Figure 57): bilateral, hub, and hierarchical.

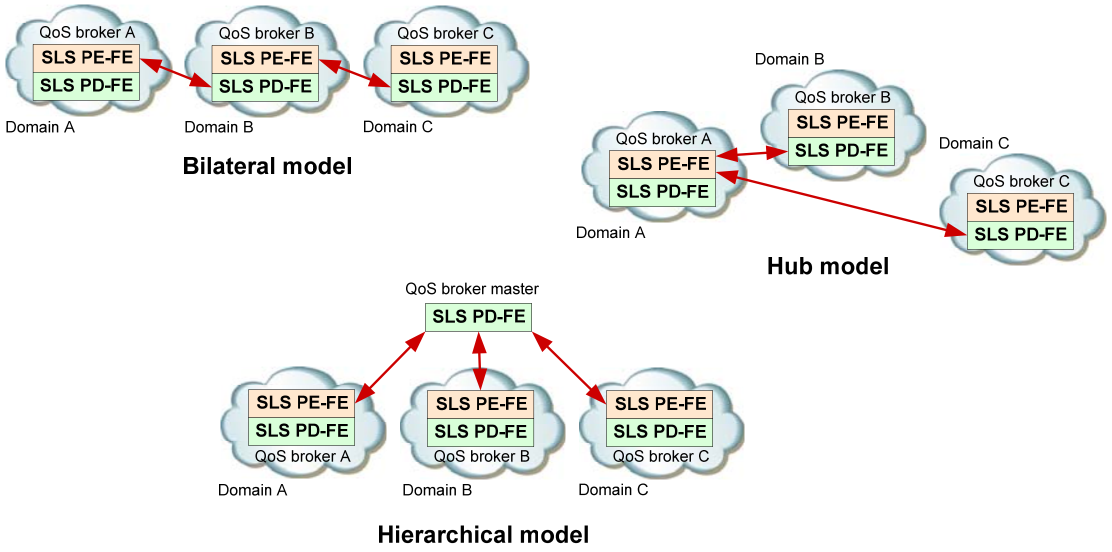


Figure 57: COPS-SLS negotiation models

In the case of an inter-domain negotiation, where the service has to pass through multiple domains, the resource request is propagated from PD-FE to PD-FE throughout the path until it reaches the last PD-FE where is confirmed. Then this confirmation is returns back to the initial PD-FE on the same path. Finally there is an end-to-end acknowledgement signalling. It has the same behaviour as SIBBSv6 since a request is propagated from one bandwidth broker to the other in each domain of the data path, except for the final acknowledgement, which provides more guarantees to the whole system. The message sequence chart of the negotiation process between 3 domains is illustarted in Figure 58. When the service has to pass through multiple domains (inter-domain negotiation), the resource request is propagated from QoSB A (SLS PE-FE) to QoSB B (SLS PD-FE) throughout the path until it reaches the last QoSB where it is confirmed. Then this confirmation is sent back to the initial QoSB on the same path. Finally there is an end-to-end acknowledgement signalling. As it is a client-server architecture, QoSB will take on the dual role of PDP (SLS PD-FE) and PEP (SLS PE-FE). At each intermediate domain, the SLS PD-FE retrieves information from the policy repository and decides the service class it can offer the communication. The SLS PE-FE of these domains selects the predefined SLS for the next request.

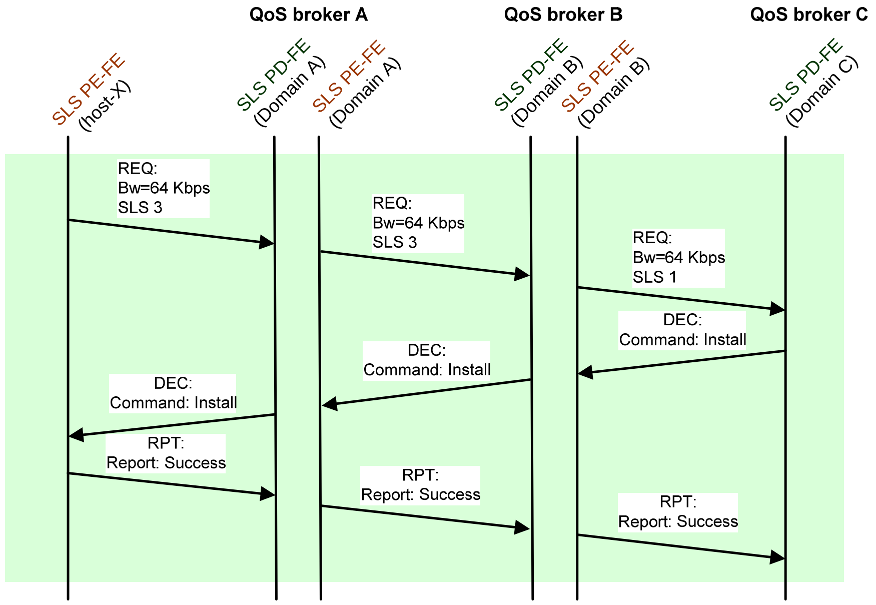


Figure 58: Sequential message diagram for end-to-end negotiation with three domains

COPS-SLS was developed in order to potentially provide to the network management a homogeneous environment by the use of COPS in both the negotiation and service deployment processes. The same way as COPS-PR, COPS-SLS also uses a PIB to store the network policies representing the SLS information. It was thought this would enhance the re-usability level of the protocol. If an optimization analysis is carried out, the election of the COPS-SLS protocol may also result in an internal resource optimization of the QoSB and may facilitate the interoperability between the resource assignation protocols of different interfaces. This chapter carries out such analysis which is shown in section 4.6.

4.5. Proposed architecture for end-to-end QoS control

An implementation of a QoS Broker (QoSB) has been proposed to create centralized management of QoS resources according to the NGN-GSI architecture guidelines, which some authors have denominated Bandwidth Manager (MSF) or Bearer Resource Manager (ITU-T) [120]. This QoS Broker is the evolution of the BBv6 developed in the previous chapter. The proposed QoSB assumes the PD-FE and TRC-FE functions (RACF). Nevertheless, its functions are not limited to those defined in RACF as it also takes on the network policy delivery to domain nodes which will probably be defined in Rn interface.

In the following examples a user A establishes a call with user B. The first step is to stablish a SIP session with the SCCF. When user B agrees to answer the call (or the multimedia communication), the SCCF requests to the RACF if there exist enough QoS resources in order to accept the communication. Whenever a service flow resource is admitted in a given domain, the QoSB notifies the flow identification, the route and the QoS attributes to the edge nodes. These nodes will identify, classify, mark, policy and encapsulate the flow packets with the QoS information specified by the QoSB. If the service flow needs to pass through multiples domains, it will be the edge nodes (or gateways) which will interconnect the different domains through the specified inter-domain SLAs from the signalling aspect. Figures 59 and 60 show an example of end-to-end signalling proceeding for a single operator managing three domains for the two cases, the push mode and the pull mode respectively.

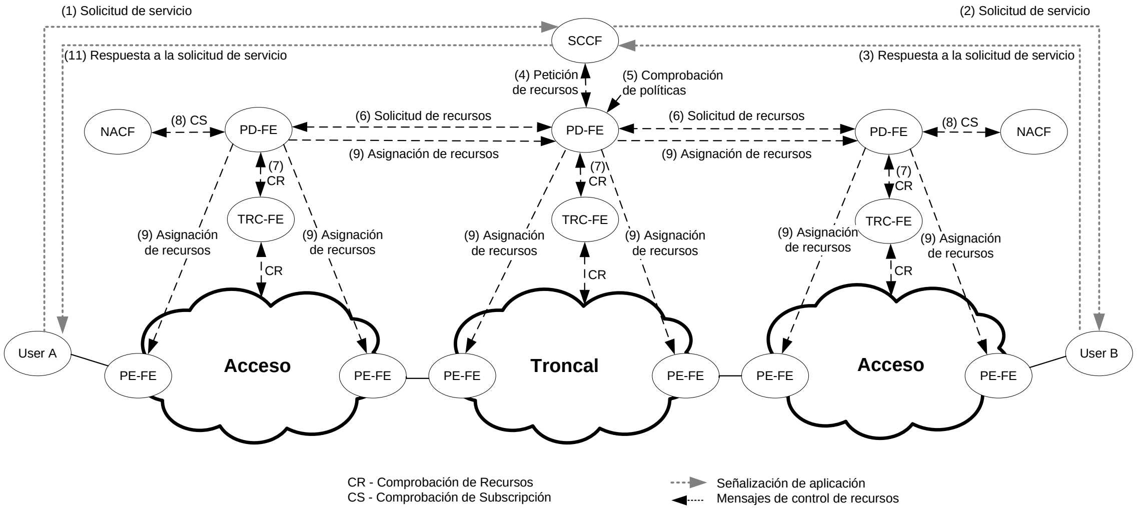

(1) Solicitud de servicio
(2) Solicitud de servicio
SCCF
(11) Respuesta a la solicitud de servicio
(3) Respuesta a la solicitud de servicio
(4) Petición de recursos
(5) Comprobación de políticas
NACF
(8) CS
PD-FE
(6) Solicitud de recursos
PD-FE
(6) Solicitud de recursos
PD-FE
(8) CS
NACF
(9) Asignación de recursos
(9) Asignación de recursos
(7) CR
(7) CR
(7) CR
TRC-FE
TRC-FE
TRC-FE
(9) Asignación de recursos
(9) Asignación de recursos
(9) Asignación de recursos
(9) Asignación de recursos
(9) Asignación de recursos
(9) Asignación de recursos
CR
CR
CR
User A
PE-FE
Acceso
PE-FE
PE-FE
Troncal
PE-FE
PE-FE
Acceso
PE-FE
User B
CR - Comprobación de Recursos
CS - Comprobación de Subscripción
Señalización de aplicación
Mensajes de control de recursos

Figure 59: Example of push-mode end-to-end signalling with three domains

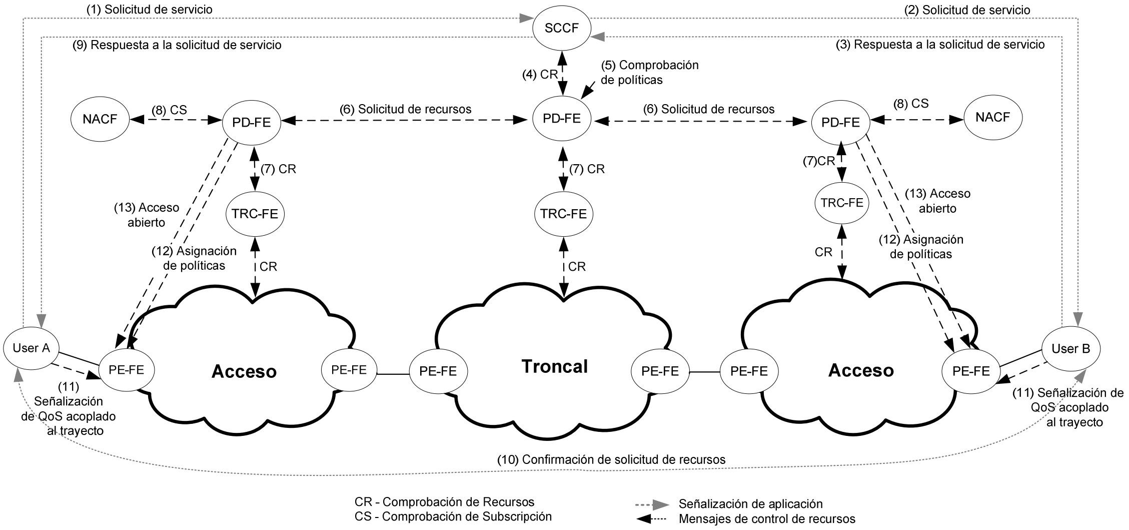

(1) Solicitud de servicio
(2) Solicitud de servicio
SCCF
(9) Respuesta a la solicitud de servicio
(3) Respuesta a la solicitud de servicio
(4) CR
(5) Comprobación de políticas
NACF
(8) CS
PD-FE
(6) Solicitud de recursos
PD-FE
(6) Solicitud de recursos
PD-FE
(8) CS
NACF
(7) CR
(7) CR
(7)CR
TRC-FE
TRC-FE
TRC-FE
(13) Acceso abierto
(13) Acceso abierto
(12) Asignación de políticas
(12) Asignación de políticas
CR
CR
CR
User A
PE-FE
Acceso
PE-FE
PE-FE
Troncal
PE-FE
PE-FE
Acceso
PE-FE
User B
(11) Señalización de QoS acoplado al trayecto
(11) Señalización de QoS acoplado al trayecto
(10) Confirmación de solicitud de recursos
CR - Comprobación de Recursos
CS - Comprobación de Subscripción
Señalización de aplicación
Mensajes de control de recursos

Figure 60: Example of pull-mode end-to-end signalling with three domains

4.5.1. First proposal with COPS-SLS

An end-to-end QoS management signalling proposal is presented in order to meet the specifications of the NGN-GSI architecture. This proposal was designed in 2008 according to the early ITU-T specifications for the Ri interface. These specifications proposed dynamic service level negotiation protocols. The study from Sarangan et al. [181] helped in the decision of using COPS-SLS as an inter-domain protocol as discussed in the previous section.

Figure 61 summarizes the protocols proposed which conform with ITU-T specifications, where signalling between a QoS Broker (QoSB) and the nodes of the same domain will be carried out with the COPS-PR protocol, signalling between adjacent QoSB will be carried out with the COPS-SLS protocol and resource requests received by the SCCF inside its own domain will be subject to the DIAMETER protocol.

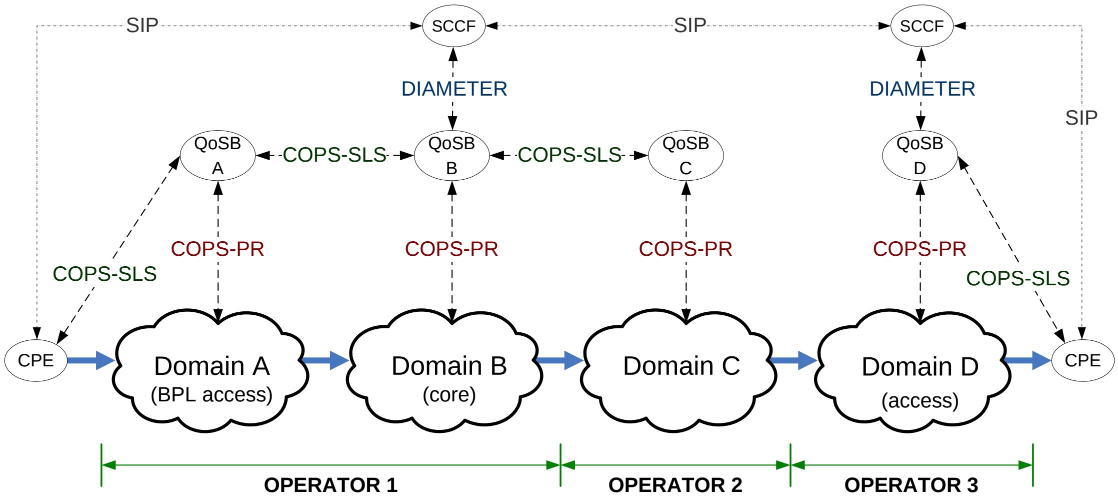


Figure 61: Proposed architecture for end-to-end signalling

4.5.2. Hybrid proposal with COPS-SLS and DIAMETER

By the end of 2008 the IETF created the Diameter Maintenance and Extensions WG (DIME WG) [94] in order to further develop the Diameter protocol. One of the work focuses is on extensions to Diameter to support QoS information to be authorized and provisioned in the DIAMETER protocol was natively designed for authentication, authorization and accounting (AAA) deployments. The idea is to develop a communication protocol able to dynamically negotiate service levels of QoS in a non trusted enviroment where security is a must. This protocol is being developed for inter-domain communication between two operators who do not trust each other. In August 2009 released the RFC 5624 about Quality of Service Parameters for Usage with Diameter [128] and in February 2010 released the RFC 5777 about Traffic Classification and Quality of Service (QoS) Attributes for Diameter [129] which have been rapidly pushed to the Standard's track as a Proposed Standard.

The second proposal for end-to-end signalling protocols is shown in Figure 62 which conform with the latest approved specifications of the ITU-T NGN-GSI architecture. COPS-SLS conforms early specifications for the Ri interface. COPS-SLS is proposed in as an alternative to DIAMETER whenever the relationship between QoSB is completely trusted (same operator). COPS-SLS presents some advantages over the planned DIAMETER protocol with QoS extensions such simplicity, direct complementation with COPS-PR and end-user support. The DIAMETER protocol will be the ITU-T proposed protocol for the Ri interface for general interaction within RACF entities, even though is pre-published at the time of writing this thesis and the protocol is still being streamlined.

The COPS-PR protocol is proposed given the fact that is the common protocol for RACF interfaces towards the transport layer, it manages policy delivery on a native level and that it is already supported by main core routers. In Figure 62 the thick blue arrows represent user multimedia data and the slash-dot arrows represent the session signalling needed to provide end-to-end QoS. This protocol architecture may also be used to provide security but this is out of the scope of this thesis.

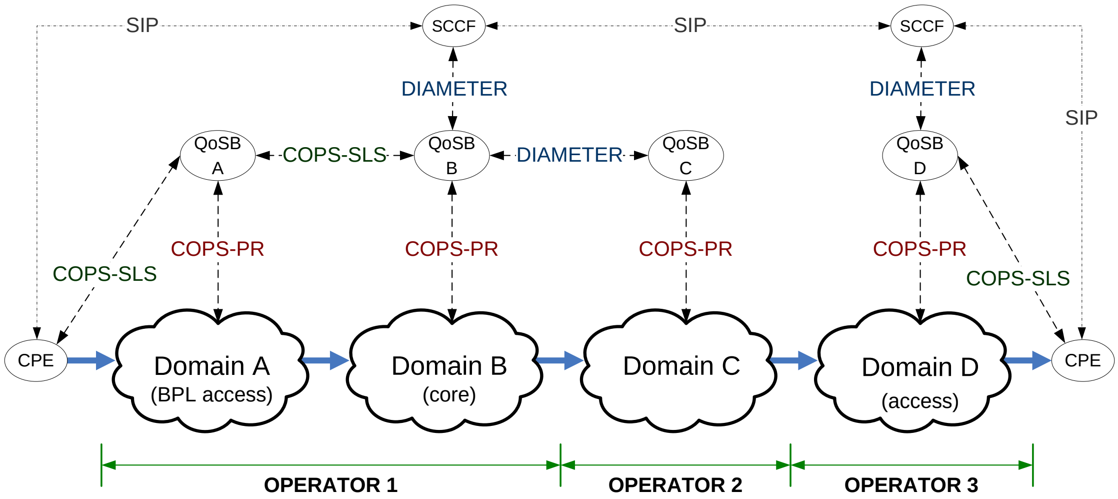


Figure 62: Proposed architecture for end-to-end signalling

4.6. Optimizing the COPS-PR and COPS-SLS Interaction

The purpose of this section is to propose an optimization of the usage of the COPS protocol in the QoS Broker. It is focused specifically on the Policy Information Base (PIB), a database which contains information about policies. To achieve this, the mechanisms of the COPS protocols implemented in the QoS Broker and their usage in the PIB are analyzed and it is proposed a new COPS variant, which we have called COPS-E2E, in order to optimally meet ITU-T recommendations for NGNs.

Both protocols use the same COPS header, albeit for slightly different purposes. This is an important advantage because in order to optimize, the re-use of as much code as possible is an added value factor. However the main challenge when trying to optimize the simultaneous use of COPS-PR and COPS-SLS lies in the Policy Information Base (PIB).

A PIB is a virtual information store where the SLSs are stored in form of policies. These policies are introduced either by the administrator or through negotiation with another domain and are stored in the domain's policy repository where the PDP can administer the policy management of the domain. The first advantage of using PIBs is that the protocol becomes independent of the policy information carried. This characteristic allows it to avoid revisiting the protocol every time new information needs to be carried by the protocol. The second advantage is that the classes defined in a PIB are reusable or extendable in other PIBs.

The Structure of Policy Provisioning Information (SPPI) is described in [16] as a structure for specifying policy information that can then be transmitted to a network device in order to configure its policy. This SPPI defines the adapted subset of SNMP's Structure of Management Information (SMI) which uses an adapted subset of the ASN.1 data definition language and is used to write PIB modules. The model underlying this structure is one of well-defined provisioning classes (PRCs) and instances of these classes (PRIs) residing in PIB.

PRCs are objects of a PIB which can be defined as unique attributes of a policy. A PRC is identified by a unique Object Identifier (OID) registered by ICANN. Each PRC saves a set of the same data type and is composed by Policy Rule Instances (PRIs) which are the values representing the network policies that may be installed in a given device. Each PRI is identified by a Provisioning Instance Identifier (PRID).

4.6.1. COPS-PR protocol PIB

COPS-PR uses the PIB defined for DiffServ [17] to configure and manage the policies in a domain. Each interface of a managed device (PEP) shares a PIB with the PDP which administers it. When a new policy (SLS) is entered by the administrator the PDP sends it to the PEP which installs it in its PIB. Therefore both PIBs (PEP and PDP) are always synchronized.

There are some differences between the PIB of the edge routers and that of the core routers. All core routers share the same PIB as they must be as simple as possible. In a DiffServ domain their function is to queue packets by prioritizing them according to their DSCP mark. The PIBs of the edge routers will all be unique as they must apply admission policies. An example of each PIB can be found in Chapter 3.

Therefore in COPS-PR each client supports a non-overlapping and independent set of PIB modules. A framework for PIB in [118] defines a set of PRovisioning Classes (PRCs) and textual conventions that are common to all clients whose provision policy uses COPS-PR protocol. The COPS-PR PIB is formed by 3 main classes: Capability Class, Policy Class and Policy PIB Conformance [24]. Using the objects defined [179], the Capability Class detects the limitations of the devices and provides the PDP with an intelligent configuration. The Policy Class is formed by the elements which define a SLS (Data Path, Classifiers, Meters, Actions, Algorithmic Droppers, Queues and Schedulers).

The PIB of the Edge routers will be composed of the aforementioned elements, while the PIB of the core routers will mainly use the classifiers to distinguish DSCP marked traffic and the Queuing Elements (Algorithmic Droppers, Queues & Schedulers) which are inter-related in their use of queuing techniques.

4.6.2. COPS-SLS protocol PIB

The COPS-SLS PIB is oriented to policy negotiation and will synchronize the PIB at the SLS-PDP (in QoSB A) and the PIB at the SLS-PEP (in QoSB B).

This PIB is more concrete than COPS-PR's PIB and it is composed of Capability Classes, Policy Classes, Parameter Classes and Report Classes. The Capability Class represents the set of configuration information delivered by SLS-PEP to SLS-PDP in the request message. This information indicates the SLS-PEP's capacity to use the different negotiation modes through pre-defined SLSs. The Policy and Parameter classes have the negotiated SLS specifications and are used in both the configuration and negotiation phases. The Report class is used to indicate the success rate of communication through the report messages.

4.6.3. Protocols Interaction

It has been decided to analyse the protocols interaction by using the COPS-PR PIB as a base and update it with the COPS-SLS requirements because the COPS-PR protocol is in the Standard's Track while COPS-SLS is still only a draft proposal. Moreover the PIB of this latter is smaller.

CAPABILITY	COPS-PR Capability Classes	Device Capabilities
	COPS-SLS Capability Classe	Negotiation Capabilities
POLICY	COPS-PR Policy Classes	Data Path, Classifiers, Meters, Action & Queuing Elements.
	COPS-SLS Policy Classes	Negotiation Table, SLS Table
	COPS-SLS Parameter Classes	Instances for the Policy Classes
CONFORMANCE	COPS-PR Policy PIB Conformance	Summary of COPS-PR PIB
	COPS-SLS Policy PIB Conformance	Summary of COPS-SLS PIB
REPORT	COPS-SLS Report Classes	Report messages indicate communication success

Table 14: Unified PIB for COPS-PR and COPS-SLS

It has been defined four main classes of the unified PIB: Capability Classes, Policy Classes, Report Classes and Conformance Classes, formed by the respective original PIB classes as shown in Table 14.

Our extensive analysis of both PIBs has shown that only a few parts of them may be re-used. This is because even though they apparently have similar functions, they have different purposes. None of the PIB PRCs in the Capability classes can be re-used, the Report Classes are unique for COPS-SLS and the Conformance Classes are just information summaries. In fact, the only part which can be re-used in the policy fusion of both PIBs is the part of the Policy Classes in which the Edge Router faces the peer domain where the negotiation takes place. To perform this, a definition and appendage of the PRCs of the COPS-SLS PIB to the COPS-PR PIB was required.

As previously stated, both protocols use the same COPS header. Therefore all the code referring to the basic COPS messages can be unified so that each protocol has only to re-use it with its own objects.

In order to automate interaction between both protocols it has been implemented a control mechanism over the COPS-SLS negotiated policies in the QoSB that convert them into COPS-PR format. This mechanism configures the attributes to be installed in the classifiers, creates a new automated SLS for COPS-PR with COPS-SLS information (Traffic identification, Scope, QoS Requirements), saves them into the proper edge router PIB and proceeds to update it. Specifically, information from "Traffic Identification" and "Scope" is used in the Classifier PRC of the COPS-PR PIB and "QoS Requirements" is used in the Datapath elements. Moreover, another algorithm checks for duplicities or inconsistencies between the new SLS and old ones.

Both protocols were created to be complementary and therefore it could be proposed a new protocol joining their strengths. This new protocol proposal could be named COPS-E2E protocol.

4.6.4. COPS-E2E protocol proposal

The COPS End-To-End protocol (COPS-E2E) is a new theoretical proposal of a new protocol based on COPS to deliver both intra-domain policies and negotiating inter-domain policies. This protocol will guarantee an end-to-end QoS signalling management between two CPEs (or End Devices) as RSVP would do but with the advantages of the provisioning model.

Therefore, while COPS-E2E benefits from the strengths of the standardized COPS-PR protocol, the interaction of COPS-SLS enables it to simultaneously solve some of the weaknesses of COPS-PR such as its ineffective implementation in push operating mode or its inability to deliver inter-domain policies.

To append the inter-domain policy delivery inability this new protocol will use the COPS-SLS two-phase concept: Configuration and Negotiation.

In order to manage all the devices in this new end-to-end environment the role of the CPE must be defined. Thus, we will differentiate between three instances instead of two: PD-FE, PE-FE and ED (End Device). The ED will begin communication by contacting the PDP (in the QoSB) as defined in COPS-SLS. If another domain has to be contacted, then the QoSB will act as a PE-FE for the new inter-domain communication. This new functionality will thus solve the inefficiency of the COPS-PR push operating mode.

The objects to be used with COPS-E2E must be redefined while the unified PIB model proposed in this thesis may be used to determine the new PIB.

Furthermore, it is also proposed the incorporation of the COPS-feedback concept (RFC-3483) [170] for inter-domain communication. The original COPS-feedback was developed as an extension of COPS-PR and provides information on the status of a PE-FE which enables the PD-FE to carry out a more efficient management. The incorporation of this new functionality into COPS-E2E would enable a SLS PE-FE to know the bandwidth limitations which it might be conceded whenever it wanted to make a SLS request to access a new domain, thus avoiding policy renegotiation messages. This functionality can be implemented by using the predefined SLS of COPS-SLS and dynamically updating them at determined intervals during communication between two QoS Brokers.

4.7. Evaluation and results

The validation of the proposed architecture has been carried out over the testbed implemented in the previous chapter (Figure 41). The main goal of this test has been to achieve an overall system evaluation at the transport stratum, leaving more extensive performance and efficiency tests out of the scope of this thesis. More specifically the test presented has been undertaken with type-1 users and therefore in push mode.

The testbed has three domains with GNU/Linux nodes. All of them have PE-FE modules supporting COPS-PR for IPv4 and IPv6 networks, even though ITU-T is still developing the signalling adaptation of the NGNs to IPv6 (recommendation ITU-T Y.2051) [118]. All the QoSBv6 have PDP modules supporting COPS-PR and COPS-SLS, and PEP modules for COPS-SLS. In this scenario the COPS-PR protocol is used to manage the intra-domain policies and resources with the GNU/Linux routers (Rw, Rc and Rn interfaces) and the COPS-SLS protocol was used to dynamically negotiate the inter-domain policies among the centralized devices of each domain (Ri interface). Both protocols operate successfully but have been implemented with independent codes and PIBs for these concrete tests.

It has been decided to use the same testing environment as in the previous chapter in order to contrast the results. Therefore the implemented policies for inter and intra domain communication have been the used in section 3.7.3 of the previous chapter. The same SLA between the two clients (Client-X and Client-Y) and Domain A have also been used, as well as the same traffic sources (6 UDP sources marked as EF, AF21, AF41 and BE), and the same congested link in 2001:720:818:4001:2::0/80.

Table 15 presents the average results of 5 tests. In this tests once the policies are applied, the system stops dropping EF packets, it reduces the AF41 packets drop to 3.98 % and the AF21 packets drop to 5.62 % and increases the BE packets drop to 34.9 % (average of both sinks). These results are coherent with the policy applied since the EF queue has absolute priority and AF41 has a lower discarding probability than AF21. These results are very similar to the ones obtained in Section 3.7.4 with the SIBBSv6 protocol. The difference is the dynamic SLS negotiation carried out by the COPS-SLS protocol. These results are presented as a contribution of this thesis.

	Sink-1		Sink-2	
	Before policies are applied	**After policies are applied**	**Before policies are applied**	**After policies are applied**
EF	12.28 %	0 %	15.91 %	0 %
AF41	19.71 %	4.98 %	-	-
AF21	-	-	20.35 %	5.62 %
BE	15.12 %	34.37 %	16.26 %	35.09%

Table 15: Percentage of dropped packets using COPS-SLS for inter-domain communication

The results show the practical viability and operatability of the architecture even though no detailed study of the performance has been carried out. Therefore there is no information about its real efficiency and it cannot be compared it to other dynamic service negotiation protocols.

4.8. Chapter summary

A signalling architecture proposal for end-to-end QoS Management according to the ITU-T specifications for NGN Networks has been proposed, implemented and evaluated. This architecture is managed by a QoS broker which implements the RACF functions and is an evolution of the previously developed BBv6. The ITU-T is still developing the NGN-GSI architecture and some of the defined interfaces are only basically profiled. In fact, the ITU-T will shortly propose a variation of the DIAMETER protocol as an inter-domain communication protocol even though it is yet to be developed.

This study has been carried out in two phases. In 2008 the research study of the state the art, the first signalling proposal, the QoS broker implementation and the system evaluations were carried out. In 2010 the information has been updated and a new signalling proposal has been proposed according to the current state of the NGN-GSI specifications.

The QoS broker provides a device to centrally manage DiffServ domains which permits the delivery of network policies through the COPS-PR protocol, a standardized intra-domain communication protocol that optimizes policy delivery, to IPv4 and/or IPv6 domain nodes. The implementation of the inter-domain signalling among peer QoS brokers in different domains has been done with the COPS-SLS protocol. This protocol is able to dynamically negotiate QoS policies, has a very scalable architecture and is compatible with COPS-PR. The QoS broker has been programmed with JAVA for multiplatform support and it is open source. Its modular design allows specialized modules for each of the tasks it is in charge of and permits adaptability for new technologies and protocols in the future.

The optimization possibilities of interoperating COPS-PR and COPS-SLS protocols in order to improve efficiency by focusing on their PIBs have been analyzed. This analysis has shown that only very few parts of both PIBs could be re-used in a combined PIB. Finally the base for a new protocol named COPS-E2E has been sketched out. Within this protocol the fusion COPS-PR and COPS-SLS was proposed as we considered them complementary and able to solve their mutual deficiencies.

The proposal has been successfully validated at NGN-GSI architecture transport stratum level through the evaluation with a three-IPv6 DiffServ domain testbed with three QoS brokers managing each of these domains. The results of the tests show how the QoS brokers successfully negotiate policies with other peer QoS brokers, deliver QoS policies with the COPS-PR protocol to the DiffServ domain nodes and how they are correctly mapped so that the GNU/Linux OS can successfully prioritize the traffic.

These contributions to the thesis have been presented in [205], [206] and [210].

5. QOS CONTROL IN NEXT GENERATION BROADBAND POWER LINE NETWORKS

5.1. Introduction

With the increase in multimedia communications, the QoS control has become especially critical in the Next Generation Access networks (NGAs) which are usually the traffic bottleneck in end-to-end flows [187]. Improvements in the features of the broadband Power Line Communication (PLC) networks have led to their integration into NGAs. Nowadays, these networks provide high-speed symmetric and cost-effective alternative "last mile" broadband access.

The ITU's NGN-GSI generic architecture takes into account the access domains, but not all of these domains are ready to manage QoS under the necessary end-to-end requirements. The broadband PLC is one of the technologies which is about to be standardized. PLC networks are very unpredictable and therefore QoS has to be a major consideration. One of the aims of the NGN–GSI architecture is to provide seamless end-to-end QoS over heterogeneous networks for the growing multimedia traffic flows [119] but inter-domain QoS is one of the least analyzed aspects in PLC networks, even though international projects such as 6POWER [232] or OPERA [160] have made proposals. Moreover, the draft standard proposal IEEE P1901 [90] has improved QoS but mainly from the intra-domain aspect.

The future of electrical utilities goes hand in hand with the Smart Grid and its advantages. Smart Grids will save energy and can cope better with the unpredictable supply from renewable energies [44][50]. At present, utilities have to be prepared to face the increased needs of its telecommunications infrastructure. In fact, one of the main challenges of Smart Grids is to redesign its communications network architecture. The current utility grid scheme is relatively easy to operate but the Smart Grid is much more complex. Its architecture is based on a decentralized scheme with elements logically identified but not geographically located. The coming Smart Grids will manage lots of real-time information through a data network and they will collect information from established IED (Intelligent Electronic Devices) for control purposes. This kind of data network is not exempt from the growing need of QoS. Smart Grids are expected to face a drastic increase of information demand, communication and various data such as voice, data, image, video and multimedia communications, which will have to be accessed anywhere at any time. The Smart Grid is just beginning to be outlined. Since these networks will have to manage huge quantities of real-time traffic over a heterogeneous network which may consist of composed wireless and PLC systems, the QoS parameters must be internally mapped between Smart Grid communication technologies and externally mapped within other technologies in order to provide end-to-end QoS.

5.1.1. Objectives

The first objective of this chapter is to propose a QoS architecture for OPERA PLC access networks. It is proposed the development of a QoS Broker based architecture for the PLC access network in the framework of the ITU's NGN-GSI architecture. This architecture, which has been developed in the FP6-IST-OPERA project, must permit better QoS management in the PLC domain and must facilitate the interoperability with other technology domains. This architecture must control the layer-2 QoS of the PLC access network in a centralized way, so that it can make use of the knowledge of all the connections in order to provide a better QoS, and it must communicate with neighbouring QoS Broker controlled domains. Therefore the PLC Qos Broker architecture must be integrated into the core network in order to attain end-to-end QoS.

The second objective is to develop a similar architecture but for the IEEE P1901 draft proposal of a standardized PLC architecture. This proposal has improved intra-domain QoS, but does not satisfy neither end-to-end QoS control needs nor does it satisfy the centralized needs of the ITU-T's NGN-GSI architecture. This solution needs to enhance QoS in the PLC domain by providing high level QoS management and take decisions based on current network data traffic flows, existing QoS policies and customer SLA, as long as facilitating interoperability with other technology domains. The key feature of these end-to-end QoS control proposals is the mapping of QoS parameters between PLC and other technologies.

Finally, the last objective is to analyse the QoS control necessities in the Smart Grids. Although a detailed solution cannot be proposed as it is still at its first steps of definition, an initial approach can be provided. This solution is basically focused on the use of the QoS broker as QoS mapper between technologies which form part of the Smart Grid, since it is a heterogeneous network which consists of wireless and PLC nodes. This mapping scheme should be incorporated into the resulting architecture.

5.2. Broadband PLC technology

The Power Line Communication (PLC) networks are commonly divided into narrowband

and broadband. Narrowband PLC is currently being used for electric company communications, meter reading and home automation. Narrowband PLC usually uses frequencies of up to 150 KHz in Europe (450 KHz in the USA) and delivers bitrates from 2 Kbps to 64 Kbps (PRIME project with coding) [166]. BPL, by contrast, provides higher bandwidth and can be used in in-home LANs and in access networks as well as for the purpose of the Smart Grid. Usual bandwidth values of BPL are from 10 Mbps to 300 Mbps, although new systems are offering even higher bandwidths [69] and transmission frequencies up to 100 MHz [90]. This thesis it is focused on the QoS management of broadband PLC networks.

Broadband PLC networks are very unpredictable and therefore it is extremely difficult to ensure determined QoS at all the connections. Some of the problems that PLC technology has to overcome are described in [127]. These include unpredictable frequency and time dependence of impedance, attenuation and transmission characteristics, impulse and background noise and their wide variability, limited bandwidth and the harmonic interference. In the in-home scenario, the major sources of noise are from electrical appliances, which utilize the 50 Hz electric supply and which generate noise components that extend into the high frequency spectrum [130]. The variability of the channel is especially annoying for the QoS because it can make the bandwidth drop suddenly at times.

5.2.1. Standartization status

Several different proprietary broadband PLC solutions have been developed over the years with some of the better-known ones coming from the HomePlug Alliance [84], DS2 [40], Spidcom [183], Panasonic and Mitsubishi Electric. Moreover, the European Commission funded the OPERA project in order to improve PLC technology to provide broadband over the power line from 2004 to 2009. Broadband PLC technology is also undergoing two standardization processes by the ITU-T [121] and the IEEE [98].

The ITU-T is currently developing an in-home network specification, named G.hn. The Study Group 15 [123] of the ITU's Standardization Sector is working on this proposal at present and it will include power line, coaxial and phone line.

The IEEE standardization process by the P1901 working group (WG) started in 2005 and is aimed at standardizing both in-home and access networks. Four proposals contended to become the IEEE P1901 PLC standard and the OPERA project specification was one of them. Unfortunately another competitor specification has been finally chosen which has been named IEEE P1901 draft standard specification. The IEEE P1901 is a standard proposal for high speed (>100 Mbps at the physical layer) communication devices and employs transmission frequencies up to 100 MHz. This draft standard proposal is pre-published and will be officially released during the year 2010. Furthermore, the IEEE has recently been focusing on the Smart Grid by an alliance with IEEE P2030 [91], in charge of the Smart Grid interoperability issues on the Smart Grid.

This chapter focuses on both the OPERA project specification and the IEEE P1901 draft standard proposal.

5.2.2. QoS control in PLC networks

The new advances in PLC technology, particularly those referring to high speed, have positioned broadband PLC technology in the Next Generation Access Networks (NGAs) category. These broadband access networks will interconnect users through the NGN core. Therefore the provision of the necessary QoS to the new multimedia services and communications data is essential in order to avoid bottlenecks.

The QoS architecture presented in this section applies the architecture developed by ITU-T's NGN-GSI to the broadband PLC networks and is focused on OPERA proposal and the IEEE P1901 proposal. However, this QoS architecture is generic so that it can be easily adapted to other PLC architectures. In fact it was initially designed for the OPERA project in [224].

When the QoS control ITU-T's NGN-GSI architecture is applied into the PLC medium to manage the domain's QoS, the PLC devices have to support the COPS-PR protocol and/or the SNMP protocol as they will perform the transport functions and will therefore be the PE-FE and the TRC-FE. In the next sections we describe this relationship.

In order to provide seamless QoS along the end-to-end path through heterogeneous technologies, the QoS mapping must take place in the border inter-domain devices. In this case the mapping of QoS parameters must be implemented in the HE, the NTUs (IEEE P1901) and the CPEs (OPERA) devices so that each flow has suitable parameters according to each QoS specification. QoS Brokers are responsible for these mappings and they must decide whether the network has sufficient resources to handle when a new request is generated. They also have the task of the reconfiguring the mappings on demand if required. Therefore, the QoS mapping between a broadband PLC network and any of its neighbouring domains will need to be defined. Neither the IEEE 1901 nor the OPERA specifications satisfy the integration and operability of the QoS facility with any other end-to-end QoS delivery mechanism. Internetworking between different network technologies is crucial, both in Smart Grids and in PLCs, to support end-to-end QoS. The next sections provide further information on QoS mappings.

5.2.3. Related background

The 6POWER project [232] proposed an architecture for QoS managing through a centralized QoS Broker providing IPv6, QoS and PLC integration. Their solution also used COPS-PR, SIP and a QoS Broker but in a different way from the proposed by ITU-T's NGN-GSI. The main difference is the scalability of the global ITU-T proposal (vertical and horizontal) while the 6POWER proposal is a local solution for PLC domains and neither layer-3 QoS nor inter-domain communication with other QoS Brokers is provided(Figure 63).

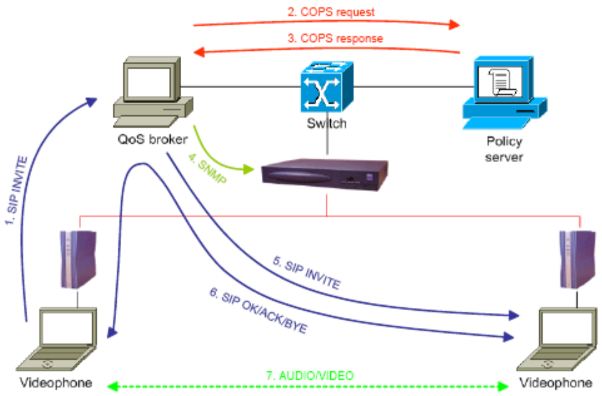


Figure 63: 6POWER architecture [232]

In the 6POWER proposal the end users connect with SIP directly to the QoS Broker and this latter connects to the policy server using COPS-PR. Once the policies a stated, they are delivered to the PLC devices using SNMP.

This project was active from 2002 to 2004 and funded under the Fifth Framework Program of the European Commission Community Research and Development Information Service (CORDIS).

5.3. OPERA specification

OPERA [160] is an access PLC architecture. It has four basic elements: the Head End (HE), the Time Division Repeater (TDR), the Frequency Division Repeater (FDR) and the Customer Premises Equipment (CPE).

The OPERA network has a tree shape architecture with the Head End in the root and the CPEs in the leaves. CPEs are the end user equipment that are located at the user's home and which connect it to the electrical network, thus connecting the end user to a PLC access network.. TDRs are repeaters which are used to relay the stream of data when signal is not strong enough. The FDR is also a repeater, but it connects two different OPERA networks. Therefore, TDR/FDR are used to extend the coverage of the network. Finally the HE is an element located at the service provider side that aggregates all the traffic from/to the CPEs and connects the PLC network to the backbone infrastructure. The HE is also responsible for controlling the access of CPEs to the network and managing the QoS. CPEs may be connected directly to the HE or through a series of Repeaters RPs. Repeaters increase the range of the PLC signal by retransmitting the signal that they receive either at a different frequency to the signal that they receive (Frequency Division) or in different time slots (Time Division). Figure 64 depicts the architecture of the OPERA network.

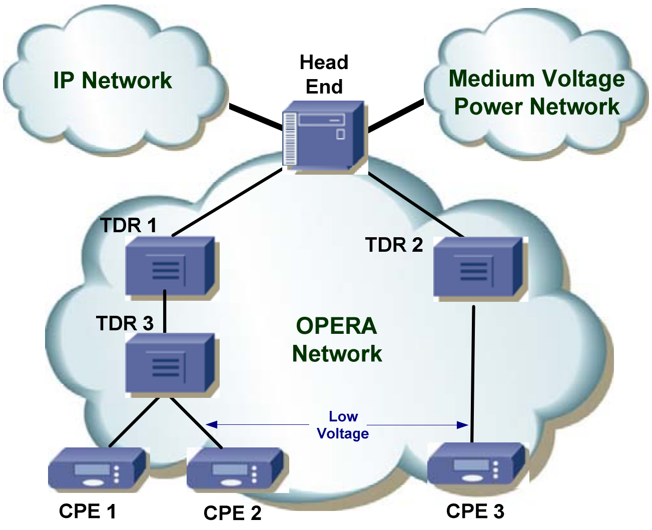


Figure 64: OPERA architecture

OPERA elements transmit in a turn-based mode using a token frame. Figure 65 illustrates how it works in a network with the HE, a TDR (REP1) and two CPEs. The QoS Broker architecture described in this chapter will manage all the QoS issues which performs the Head End and which will be explained later.

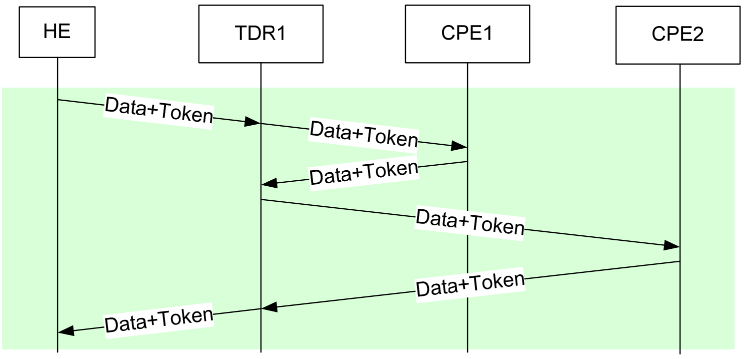


Figure 65: Token passing in OPERA

The HE controls the token and sends it to each of the TDRs it has connected. Only one token is passed between devices on the OPERA network, authorizing them to communicate. Then, the repeaters do the same with each of its children until the token arrives to the CPEs and then goes back up to the HE. The control of the time that each device poses the token is important to ensure that the network has a limited delay.

OPERA also specifies security concerns but they are out of the scope of this thesis as it is mainly focused on QoS.

5.3.1. QoS control in OPERA networks

In order to control de QoS, it is proposed to locate the QoS Broker in the IP network near the HE, being responsible for configuring the HE, the TDRs and the CPEs. Even though the OPERA network is a layer 2 network, its devices can be configured using a layer 3 protocol which makes the mapping of the ITU-T proposal possible. In this case IPv6 protocol has been recommended for the NGN. The QoS Broker will be connected to the BPL cell also with IPv6.

When applying the ITU-T's NGN-GSI architecture in an end-to-end basis, a CPE within a broadband PLC access may establish end-to-end multimedia sessions with QoS. This is even the case when the destination user is in another operator's domain but only when all the intermediate domains are controlled by QoS Brokers. The end-to-end QoS depends on the weakest QoS domain in the data path, taking into consideration of bandwidth and delay. This proposal uses a provisioning scheme using COPS-PR for the intra-domain communication and COPS-SLS and DIAMETER for the inter-domain communication as stated in the previous chapter. Therefore OPERA QoS architecture works in a push (provisioning) mode, using COPS-PR as the protocol to communicate with the transport functions.

When the QoS Broker has to allow a reservation to a destination domain which is known to the topology database, it will send a signalling message to the destination domain. If the request succeeds, the user will be able to reach the destination with a guaranteed bandwidth and QoS. The QoS Broker will be responsible for accepting the requests from users and pushing the policies to the HE, TDRs and CPEs using COPS-PR (Figure 66). As previously mentioned, these requests may come from the SCCF (SIP), the CPE (COPS-PR) or directly from the user/CPE (COPS-SLS).

If the ITU-T's NGN-GSI architecture is not available the CPE has to send a signalling message directly to the QoS Broker when it wants to make a reservation.

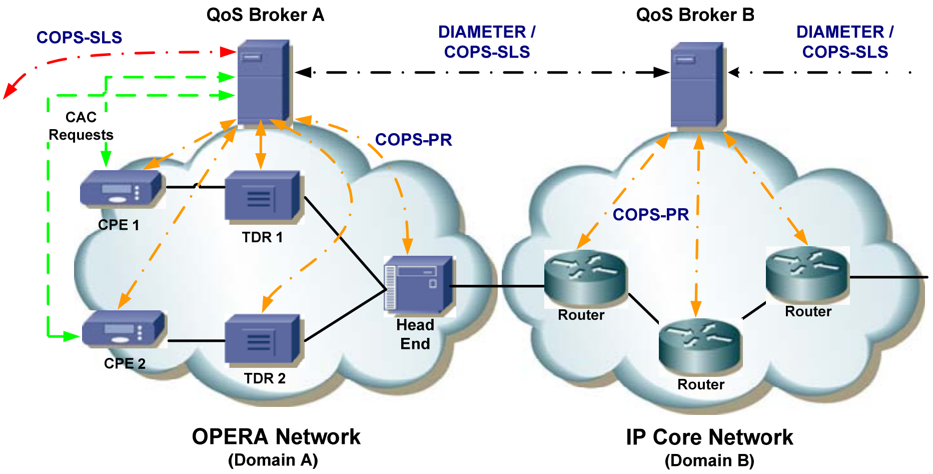


Figure 66: QoS control architecture for the OPERA network

The QoS Broker has to constantly check the network status and the link quality due to the unpredictable behaviour of the PLC networks. This is done through the TRE-FE entities which are implemented in the HE, TDRs and CPEs. The PE-FE is included in the CPEs and the HE in order to prevent unwanted data fluxes from entering the network. In Figure 67, the relation between OPERA elements and ITU's NGN-GSI elements is shown.

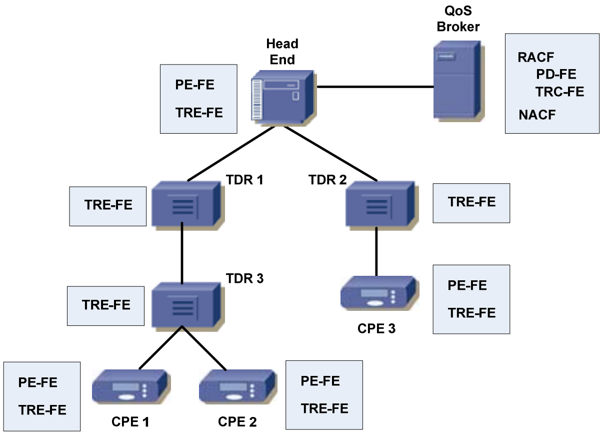


Figure 67: RACF entities mapped into the OPERA architecture

Another option would be to embed the QoS Broker in the HE, at least the RACF, leaving the NACF outsourced. This option has several advantages given that the HE is already the layer-2 QoS manager so intra-domain communications between RACF and HE (PE-FE and TRE-FE) within the Rw, Rn and Rc ITU-T's interfaces would be internal. Nevertheless, it would mean that the RACF would have to be completely standardized and included in the HE OS by device manufacturers, which is improbable in medium term.

5.3.2. OPERA QoS facility

The MAC layer with QoS support in the OPERA specification contains the required functionality to comply with the different services, and to conform to the Service Level Agreement (SLA) of each customer. The main objective of this QoS is therefore to guarantee a certain bandwidth level and latency to different users, depending on the how the available services are configured and on the type of traffic being transmitted. With the aim of dealing with this issue, a priority based system has been used.

One important issue with QoS in PLC networks is that it is very unpredictable. Because of that, it is extremely difficult to assure the QoS to all the connections. The OPERA proposal provides various mechanisms to assure the QoS of the traffic flows in the broadband PLC cell. Traffic flows are session oriented and tagged with certain service class that implies certain requirements in terms of latency and bandwith, jointly with a requirement in terms of level of assurance of the reserved resources. OPERA based systems provide up to eight different priorities, which are mapped to the eight classes of possible services available (which will be referred to as SLA). These service classes, or SLAs, are mapped to four different maximum latencies. The incoming traffic goes through a module called the Traffic Classifier, whose job is to prioritize packets based on some simple rules. Therefore each CPE will tag their flows with a priority from 0 to 7. If the bandwidth decreases suddenly, some packets will be dropped.

Four service parameters are necessary to fully describe a service class [160]:

- **Priority**: Values from 0 to 7, being 7 the highest priority. If packets have to be dropped at the transmitter because of lack of resources, then the lower priority packets shall be dropped before the higher priority packets. When there are data of different priorities addressed to the same destination, higher priority data shall be transmitted earlier than lower priority data
- **Max Subcell Access Time**: It corresponds to the maximum duration for a flow to access the channel on the subcell level. It is a latency requirement for the scheduler to be configured within any master of a PLC cell
- **Resource Reservation**: It is related to the type of claimed guarantees provided in terms of node resources and time resources. It can take the following values:
  - **Type Best Effort**: No guarantees on the node and time resources.
  - **Type ABR (Available Bit-Rate)**: Requests ask for node and time resources for which a minimum bandwidth can be guaranteed with a peak over the minimum.
  - **Type VBR (Variable Bit-Rate)**: Requests ask for node and time resources for which an average bandwidth can be guaranteed with a maximum variation around this average.
  - **Type CBR (Constant Bit-Rate)**: Requests ask for individual guarantees in terms of node and time resources. The granted node and time resources are not shared. The claimed bandwidth is guaranteed.
- **Service Reliability**: Defines if the acknowledgement mode (ACK) mode is enabled within the service class.

Service Class (SLA)	Resource Reservation	Priority	MaxSubcell Access Time / Latency (ms)	Service Reliability (ACK enabled)	Application Examples
7	Best Effort	7	240	No	Management messages
6	CBR	6	30	No	VoIP
5	VBR	5	60	Yes	Video, Games
4	VBR	4	120	Yes	Data HighPrio
3	VBR	3	120	No	Data HighPrio
2	ABR	2	120	Yes	Data HighPrio
1	Best Effort	1	240	No	Data LowPrio
0	Best Effort	0	240	No	Data LowPrio

Table 16: Example of definition of service classes [160]

A slave device requiring to transmit certain type of traffic shall first classify it in one of the eight available service classes, each one with a preprogrammed latency, bandwidth and level of assurance of the required resources. The mapping from Service Classes to the Bandwith, Latency and Type of Traffic is common and known by all nodes in the PLC cell. Table 16 presents an example of definition of service classes and could be used to implement QoS services for applications as described.

The QoS of OPERA based systems provide the following features in order to guarantee the differentiated services to every connected user:

- **Service Differentiation**: Priorizes packets with the Service Classes using a Traffic Classifier.
- **Latency Management**: It is carried out by the nodes which detect the transmitted and received priorities so that the scheduling mechanism is adapted to the type of traffic and to the specified SLAs. Thus, a node transmitting high priority traffic will receive better service than a node transmitting low priority traffic.
- **Call Admission Control:** The requirements to transmit the traffic flow are solicited to the Flow Master Node of this service class, and can be accepted and a session ID shall be assigned to it, or rejected. In the case where the maximum capacity of the channel is reached, several congestion policies can be applied by the QC and FMNs to admit, reject or drop current sessions in the broadband PLC cell.
- **Bandwidth Control**: It ensures that node's throughput is the specified value at all times. The scheduler also maintains the obtained throughput to the specified value, although sometimes it is not possible due to insufficient channel quality. Bandwidth control is performed on the transmitter side in order to maximize the efficiency of the system.
- **Congestion Management**: It is responsable of the QoS layer. Different policies can be applied in order to adapt the system to the different types of traffic. The possible policies are:
  - *Fair Congestion Management*: Performance is decreased globally for all flows.
  - Priority Congestion Management: Performance is decreased first for lower priority flows and high priority flows are not touched.
  - *Quality Service Congestion Management*: Performance is decreased first for lower priority traffics. Quality Session parameters are used to prioritize between flows with the same service class
- User Profiles: It is used to define the QoS profile of different connections in the PLC network which are special files to specify the characteristics of the connections in terms of service parameters.

A Traffic Classifier module is the recommended mechanism for the OPERA devices to identify the class of service that each specific Ethernet frame belongs to in order to handle different services and applications adequately. This module is responsible for determining the priority level of each frame according to a set of rules established by the device manufacturer, the service provider or the end-user. Typical default rules for the Service Classifier could be: decide priorities according to bits in 802.1p field, or according to bits in IPv4 TOS field, etc.

The QoS facility applies traffic shaping in two different ways: Limited bandwidth for best-effort and ABR service classes and Guaranteed bandwidth for VBR and CBR service classes. Traffic must be allowed prior to transmit and a reservation is made for this type of traffic using the CAC protocol. Channel time is assigned based on the resources reservation to meet bandwidth and latency requirements. Due to the explicit reservation of resources, nodes shall have mechanisms to allocate buffer space to support the incoming traffic. Furthermore, there are MAC primitives to give the QoS scheduler an idea of the broadband PLC topology which collect information in the Upstream direction. These parameters are the Maximum Number of Hops in the network and the Number of Downstream Nodes.

Another interesting parameter is OVLAN configuration. OVLANs are a type of VLANs used in OPERA which can improve the efficiency of the network. TDRs repeat a message to all its neighbours in the network. If the messages are tagged with a OVLAN number, TDRs relay the message only to the corresponding neighbours thus saving bandwidth.

In OPERA architecture, the QoS Broker must play the role of the RACF. It also has the responsibility for receiving the signalling messages from the CPEs. It could accept or reject new connections and converts the high level QoS parameters into layer-2 PLC QoS parameters. It is necessary to do some calculations by QoS Broker in order to map the resource reservation message to the PIB parameters. For example, it is necessary to aggregate all the bandwidths from connections with the same priority. The maximum subcell access time is calculated from the overall delay and it is necessary to use the knowledge of the topology of the OPERA network as the number of repeaters and the number of CPEs.

In order to calculate the maximum delay, the delay of all the nodes in the network has to be aggregated in one cycle (from the time the HE starts sending the token to the first repeater until the last node has returned the token). In [182] a more involved model of the delay of the OPERA network can be found.

5.3.3. PIB proposal for OPERA devices

In this section, our proposal of the PIB used by TDRs and CPEs in an OPERA network is defined. The QoS will be based on a per user reservation and each user can do one reservation. A priority based method is used in order to differentiate between flows of the same user. These flows will be bidirectional, although the upstream and downstream parameters can be different.

The PIB of OPERA is illustrated in Figure 68 using example values for better understanding. The representation of the PIB follows the same format used in section 3.7.1. As stated before, COPS-PR usually uses the PIB defined for DiffServ [24] to configure and manage the policies in a domain. Furthermore a framework for PIB in [179] defines a set of PRCs and textual conventions that are common to all clients that provision policy using COPS-PR protocol. The root element of the PIB in Figure 68 is the *opReserv*, which represents the configuration for the OPERA network of a single device and has two parameters: the device identification (*DevID*) and the role (*Role*) it plays. The role can be Head End, TDR or CPE. The *opDeviceReserv* represents the configuration of each device that has made a reservation. In the case of a CPE, there will only be one *opDeviceReserv*. In the case of a TDR or HE there will be one opDeviceReserv for each reservation made by its CPEs. The parameters of the *opDeviceReserv* are: the device identification (*DevID*), which identifies the *opReserv* it belongs to; an identifier of the device and its MAC address (*ReservID* and *DeviceMAC*). Finally, each reservation of a device contains 8 *opPriority* which specify the parameters of each priority.

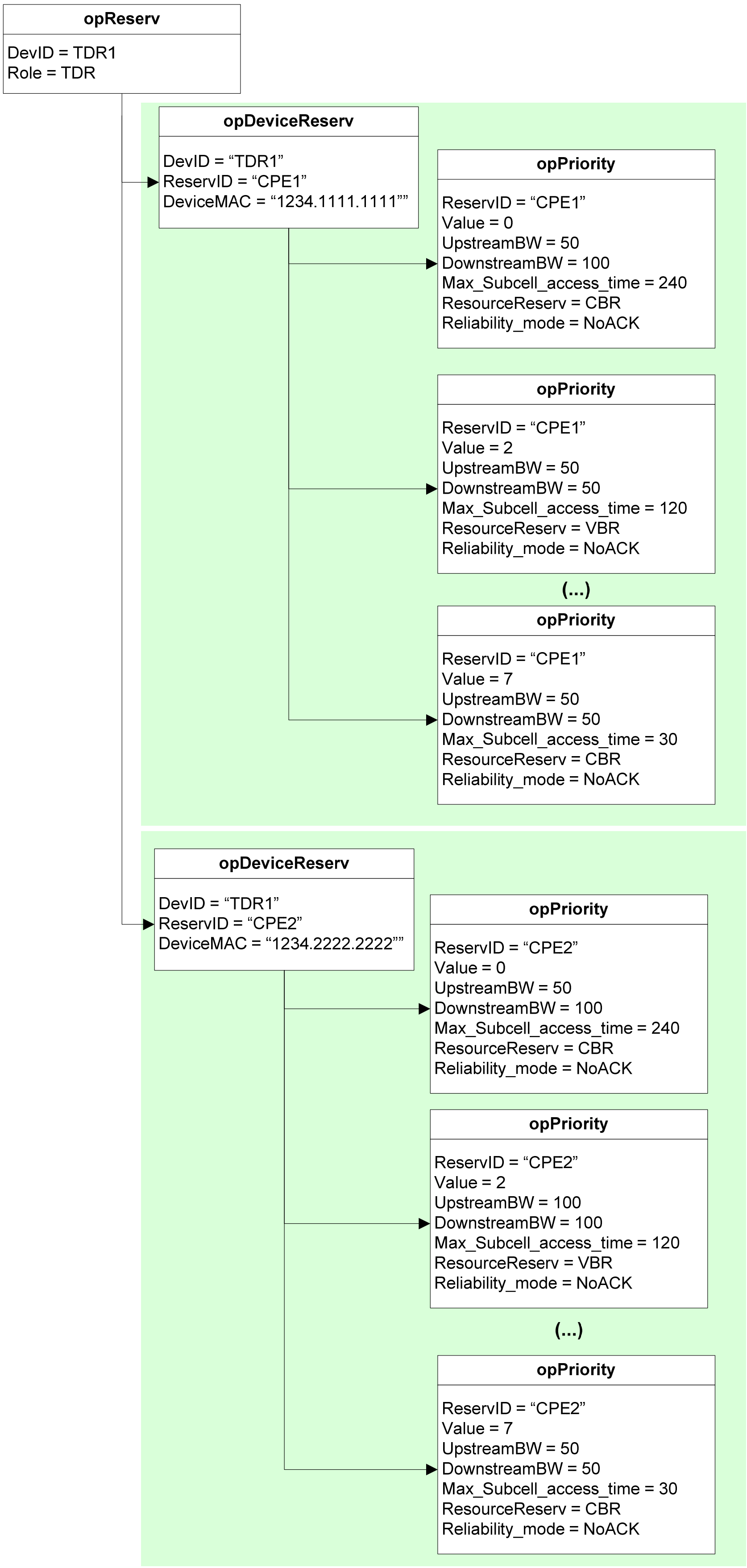


Figure 68: Example of PIB for the OPERA network

Each of the 8 priorities uses the following parameters:

- **ReservID**: It defines which of the CPEs (opDeviceReserv) the stream belongs to.
- **Value**: It defines the priority value to which the parameters are applied.
- **UpstreamBW and DownstreamBW**: The maximum upstream and downstream bandwidths denote the maximum bandwidth that can be used by a flow from a CPE with a particular priority. It could be different for upstream and downstream fluxes.
- **Max_Subcell_access_time**: It defines the maximum time a node can be transmitting that flow. This value affects both the delay and the efficiency of the network. It should be as big as possible while meeting the delay requirements.
- **ResourceReserv**: The resource reservation type parameter indicates whether the traffic is Constant Bit Rate (CBR), Variable Bit Rate (VBR), Available Bit Rate (ABR) or Best Effort (BE).
- **Reliability_mode**: The Reliability mode indicates whether the packets will be acknowledged or not.

OPERA devices define several parameters which have not been described in the PIB because they are not essential for the QoS Broker. One of them is the CoS (Class of Service) configuration. CoS configuration attaches a priority to the VLAN (IEEE 802.1Q) header and could be one of the options to be used in order to map priorities from/to the OPERA domain for improving end-to-end QoS.

5.4. IEEE P1901 specification

The draft standard proposal IEEE P1901 is formally named “Draft Standard for Broadband over Power Line Networks: Medium Access Control and Physical Layer Specifications”. The basic IEEE P1901 broadband PLC access architecture consists of a core cell, named BPL (Broadband PLC) Cell, which is connected to the in-home PLC architectures by extension sub-cells based on the AV protocol and the BPL protocol. This chapter focuses on the access BPL architecture and it will therefore only consider the BPL Cell. Figure 69 represents the architecture of the IEEE P1901 access network and its interconnections.

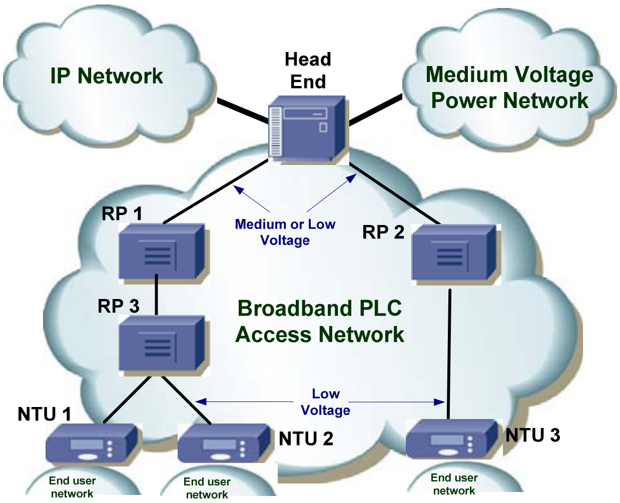


Figure 69: IEEE P1901 access architecture

The BPL Cell has three basic elements: the head end (HE), the repeater (RP) and the network termination unit (NTU). The core cell excludes customer premises equipment (CPEs) as they are part of the in-home PLC architecture. NTUs are devices that interconnect the core cell to the end users' network or devices. NTUs are similar to CPEs OPERA devices. These devices are the end user equipment that are located at the user's home and which connect it to the electrical network. The NTUs may connect the core cells to the sub-cells or bridge to other network technologies (e.g. WiFi). Each subcell has one and only one active NTU which operates as its Central Coordinator. Repeaters (RP) are devices that selectively retransmit data frames to extend the effective range and bandwidth of the BPL Cell when the signal is not strong enough. RPs are similar to TDRs OPERA devices. Repeaters may perform QoS functions. Finally the HE is an element located at the service provider side that aggregates all the traffic from/to the NTUs. There is only one active HE per cell which manages this BPL cell including the core cell and any associated sub-cells. The HE bridges the BPL cell to the backhaul network. If the HE is located at the Medium-Voltage (MV) power network, then a MV-to-LV BPL bridge is needed which may be associated to a RP (Figure 70). The HE is also responsible for controlling the access of NTUs to the network and managing the QoS.

IEEE P1901 also specifies security concerns but they are out of the scope of this thesis as it is mainly focused on QoS.

5.4.1. QoS control in IEEE P1901 networks

When applying the ITU-T NGN-GSI architecture in an end-to-end basis, a NTU within a broadband PLC access network may establish end-to-end multimedia sessions with QoS. This is even the case when the destination user is in another operator's domain.

The same way as happened with the OPERA specification, the IEEE P1901 is also a layer 2 network but its devices can be configured using a layer 3 protocol which makes the mapping of the ITU-T proposal possible. The QoS Broker will also be connected to the BPL cell with IPv6 (Figure 70).

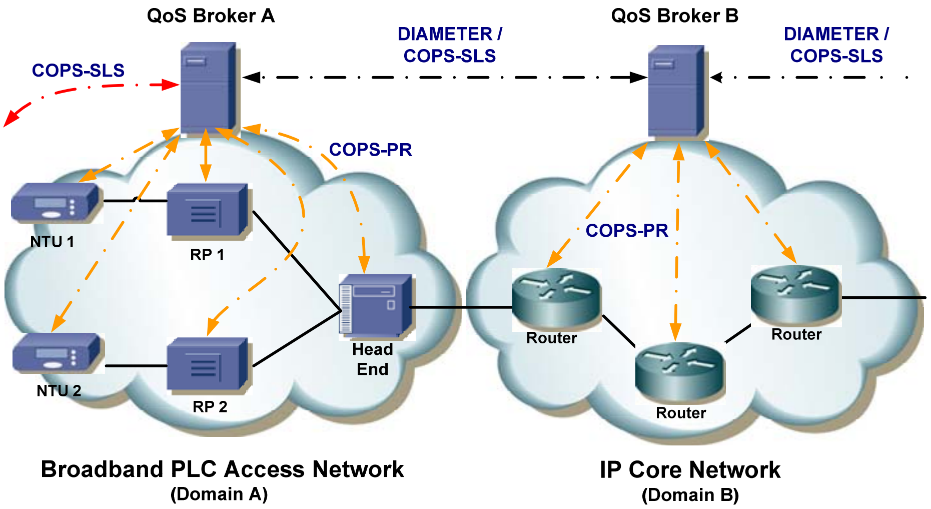


Figure 70: QoS control architecture for the IEEE P1901 network

Similarly, the proposal for IEEE P1901 also uses a provisioning scheme using COPS-PR for the BPL network. Therefore, the QoS Broker will be responsible for accepting session requests from CPE (users) and managing the policies to the HE, RPs and NTUs using COPS-PR. As previously mentioned, these requests may come from the SCCF (SIP), the NTU (COPS-PR) or directly from the user/CPE (COPS-SLS).

The QoS Broker has to constantly check the network status and the link quality due to the unpredictable behaviour of the PLC networks. This is done through the TRE-FE entities which are implemented in the HE, RPs and NTU devices. The IEEE 1901 QoS facility can be applied only to the HE and the NTUs and therefore these are also PE-FE entities. The relation between IEEE P1901 elements and ITU-T NGN-GSI elements are shown in Figure 71, which is reciprocal to the Figure 67 of the OPERA specification. As mentioned in the OPERA solution, another option for the long term is to embed the QoS Broker in the HE.

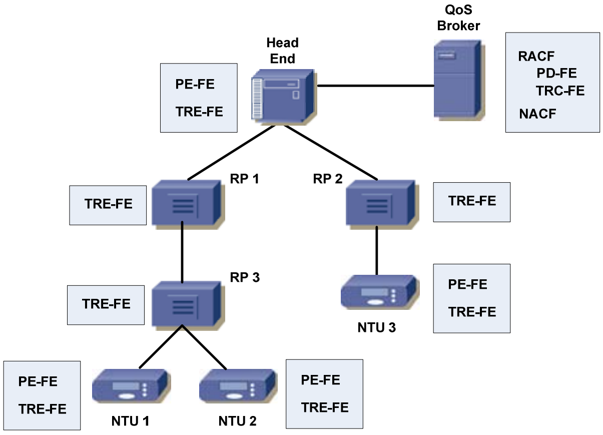


Figure 71: RACF entities mapped into the IEEE P1901 architecture

5.4.2. IEEE 1901 QoS Facility

The IEEE 1901 QoS facility is based on the IEEE 802.11e standard [92] and will be available in every associated PLC device within the BPL cell.

IEEE 1901 uses a shared medium and provides both differentiated (prioritized) and reserved (parameterized) control of access to the medium to handle data transfers with different QoS requirements. It uses a centralized QoS control function called the Hybrid Controller (HC) to determine and control the schedule for channel access. This HC uses beacon frames to notify other PLC devices of the media access schedule up to the next beacon frame. This beacon interval is a repeating synchronization process which consists of a basic media access structure issued by the HC with beacon frames. There are three blocks of time in each beacon interval: the beacon extent which is a fixed time period required to send a beacon frame, the contention free period (CFP), and the contention period (CP).

The HC operates during both the CP and the CFP and performs bandwidth management functions which include the allocation of transmission opportunities (TXOPs) to PLC devices so it has to be collocated within the HE. It provides two concurrent mechanisms for the support of applications with QoS requirements: the hybrid coordination function enhanced distributed channel access (HFC EDCA), for contention-based transfer within the CP, and the HCF controlled channel access (HFC HCCA), for contention-free transfer within the CFP.

Parameterized QoS

CFPs provide exclusive use of media by time division multiple access (TDMA). It is used for transmissions by PLC devices that require reserved bandwidth on a scheduled basis, which is to say transmissions that support low-latency, low-jitter applications with bandwidth allocations. Therefore real- time traffic such as audio and VoIP should be transmitted during CFPs. Parameterized QoS is provided through the HFC HCCA mechanism depending on the parameters associated with each data flow. It allows for the reservation of TXOPs, and therefore it is oriented to deliver layer-3 IntServ traffic. When a PLC device request for TXOP is accepted by the HC, it schedules TXOPs for both the BPL cell QoS manager (HE) and the PLC device.

The HC may provide the same level of QoS, for reserving a period of exclusive access to the medium for a PLC device, by either using a polling mechanism, to the PLC device by setting the CFP, or by using a beacon-based indication mechanism, by setting the CP. The HCCA polled access bandwidth reservation mechanism grants the transmitting PLC device with instantaneous permission to transmit for a specified length of time. With the Beacon Triggered (BT) access, after a PLC device request is accepted, the HC includes the appropriate TXOP schedule information to the reservation parameters. Therefore, a PLC device can only begin transmitting exactly at the time indicated in the beacon schedule for the allocated reservation.

Prioritized QoS

PLC devices support QoS with priority control during contention periods of access where they use CSMA/CA (Channel Sense Multiple Access / Collision Avoidance) in order to access the transmission medium. CP period is used for the “best effort” service which might include preference settings for the transmitted frames. Therefore, non realtime traffic should be transmitted during CPs. Prioritized QoS is provided through the HCF enhanced distributed channel access (EDCA) mechanism where the PLC devices use the prioritized CSMA/CA access mechanism. Therefore, EDCA delivers traffic based on differentiating user priorities and it is oriented to deliver layer 3 DiffServ traffic. Differentiation is achieved by varying the amount of time a PLC device senses the channel to be idle before back-off or transmission, and also by varying the length of the contention window to be used for the back-off, or the duration during which a PLC device may transmit after it acquires the channel. These transmissions may also be subject to certain channel access restrictions that are imposed in the form of admission control.

The EDCA mechanism support 16 levels of priority, 8 of which are user specifiable, provisioned by separate transmission queues, and selected by MAC. These 8 user priorities (UPs) are associated to 4 access categories (ACs): Background, Best Effort, Video, and Voice. For each one of these AC a set of configurable EDCA parameters are assigned:

- **CWmin**: Initial value of the contention window used to calculate the back-off time.
- **CWmax**: Maximum value of the contention window.
- **AIFSN**: Number of slots to back-off (Arbitration InterFrame Space) before back-off procedure.
- **TXOP limit**: Maximum transmission time of data frames

A PLC device waits during a back-off period before transmitting each frame. This period depends on the parameters previously defined. Depending on the length of the back-off time an AC will have a higher or lower probability of being transmitted. When a station starts a data transmission, it first senses the channel. When the channel is idle for AIFS ms, the back-off process is started. The back-off time is a random number between zero and the size of the contention window. The contention window (CW) parameter will take initially the CWmin value and will increase with every unsuccessful attempt to transmit until it reaches the value of CWmax. Once the PLC device starts transmitting, several frames might be transmitted until it reaches the TXOP limit.

The CSMA/CA contention mechanism is the same, no matter what priority the frame has. Therefore, when a PLC device has the TXOP, it must transmit the higher priority data frames first. If there is a lot of higher priority traffic to transmit, then lower priority traffic will be delayed, eventually causing significant delays.

Integration with the QoS Broker

The same way as in the OPERA proposal, the QoS Broker will be also responsible for accepting the requests from users and pushing the policies to the HE, RPs and NTUs using COPS-PR with the BPL cell QoS parameters format. As previously mentioned, these requests may come from the SCCF (SIP), the HE (COPS-PR), the NTU (COPS-PR) or directly from the user/CPE (COPS-SLS). The QoS Broker must perform some calculations before accepting any form of data flow and it will vary depending on the type of QoS requested. Once the QoS Broker has accepted data flow at the specified QoS conditions, the QoS parameters will be mapped into the Policy Information Base (PIB) parameters to be later delivered by COPS-PR to the correct devices (both QoSB and PE-FEs have a PIB). The COPS-PR protocol will transfer the new QoS policies to the devices (transferring the new PIB information) and will allow the demanded data flow once these policies have been installed.

When the end users' network uses layer-3 Intserv traffic and/or any given layer-2 parameterized QoS traffic, HFC HCCA must be used. In this case, when the NTU requests the HC for TXOP, the HC may either accept or reject the request, based on an admission control policy. The HE, where the HC is located, may forward the request to the QoS Broker in order for it to take the decision or it might take the decision on its own based on previously delivered policies. Thus, TXOP allocations and contention-free transfers of QoS traffic are based on the knowledge of instant traffic flows, the network distribution and high level QoS policies. If the request is accepted, the HC schedules TXOPs for both the BPL cell QoS manager (HE) and the NTUs. When the QoS Broker suffers a long period of loss of quality in a link segment, it ought to reconfigure QoS parameters and to tear down some accepted low priority flows.

In the case of layer-3 Diffserv traffic and/or any given layer-2 prioritized QoS traffic, the QoS Broker has to proactively deliver the policies to the edge nodes, the NTUs and the HE, based on average traffic matrix of the service contracts (SLA). It is proposed a QoS Broker which is able to change ACs contention parameters based on the knowledge of the accepted data flows and on the average network traffic. It is necessary to aggregate all the bandwidths from connections with the same priority. The maximum access time is calculated from the overall delay time and the knowledge obtained from the topology of the IEEE P1901 network as well as the number of repeaters and the number of NTUs are also key factors in the equation. It might also be useful to calculate an upper bound of the delay that frames will suffer based on the EDCA parameters. This bound can be calculated from the mathematical model proposed in [9]. Therefore, the QoS Broker will use the model to calculate the delay of the traffic flows within the IEEE P1901 network and change them to meet the end-to-end delay requirements.

In the PIB of each PLC device there will be other useful parameters to limit the size of data frames, control the size of acknowledgements or limit the transmission re-attempts, but they will not be explained in detail here.

5.4.3. PIB proposal for IEEE P1901 devices

IEEE P1901 defines two access mechanisms: HFC EDCA and HCCA. EDCA defines 8 user priorities with parameters and, therefore, it is more effectively mapped to other technologies also based on priorities. Each of these user priorities are associated to 4 Access Categories (AC). The QoS support in EDCA is carried out using those ACs and configurable parallel backoff entities.

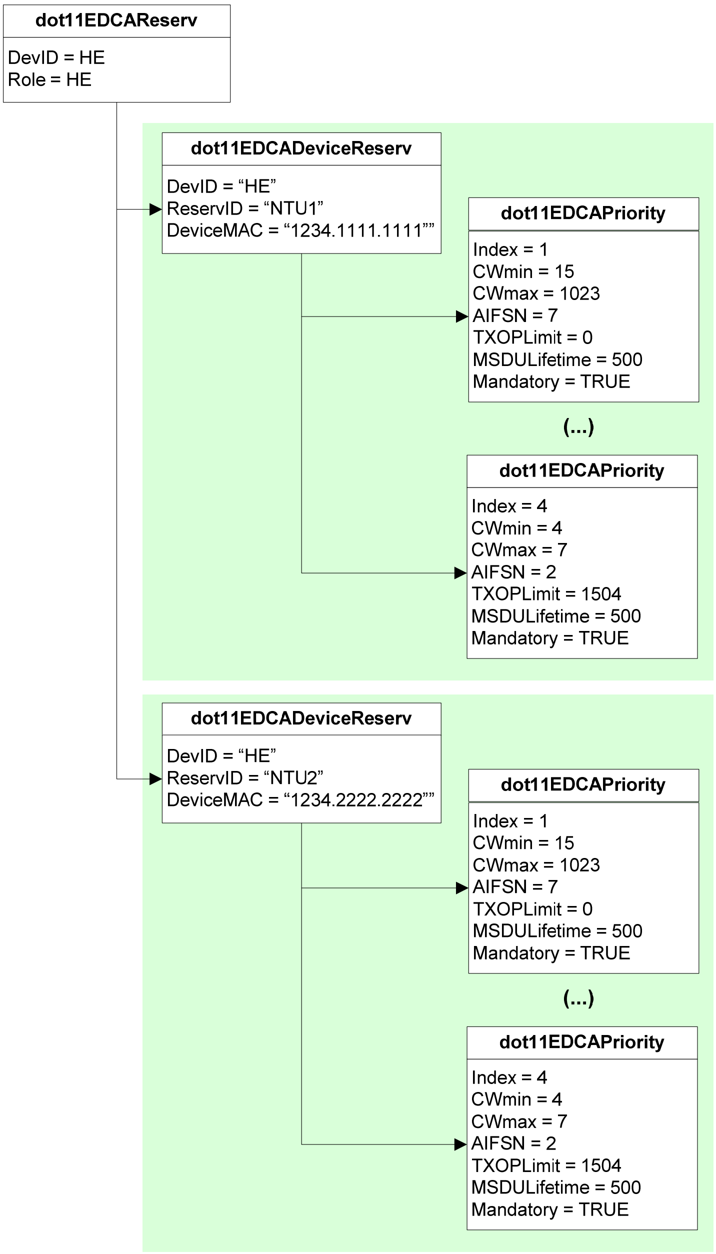


Figure 72: Example of HFC EDCA PIB for the IEEE P1901 network

It is proposed a QoS Broker able to change ACs contention parameters based on the knowledge. For each AC, an enhanced distributed channel access function (EDCAF), contends for TXOPs using a set of EDCA parameters from the EDCA Parameter Set element or from the default values for the parameters when no EDCA Parameter Set element is received from the HE, where the parameters used by the EDCAF to control its operation are defined by PIB attribute table dot11EDCAReserv. This Conceptual table for EDCA default parameter values for the registered PLC devices. For each device, this table shall contain the four entries of the EDCA parameters corresponding to four possible ACs. Index 1 corresponds to AC_BK, index 2 to AC_BE, index 3 to AC_VI, and index 4 to AC_VO. Figure 72 illustrates an example of PIB for a IEEE P1901 network with two NTUs where are defined a set of default values in the dot11EDCAPriority. The HCCA PIB is out of the scope of this thesis.

5.5. Smart Grid

The last section of this chapter discusses the QoS requirements of the Smart Grids and how to fit them into the ITU-T NGN/GSI QoS control architecture.

By and large, current power grid is defined as a system made of electrical generators, transformers, transmission and distribution lines used to deliver electricity power to final users. Monitoring and Smart Grid network control are very important features in order to provide continuity, Quality of Service (QoS) and security. Nevertheless, at the time, most of these functions are only carried out at high voltage and, sometimes, in the medium voltage grid. Actually, international organizations, governments, utilities and standardization organisms are becoming aware that the grid needs a modernization. The future Smart Grid must be distinguished by its self-healing and automation taking into account that should support thousands of clients and energy providers.

In the USA the National Institute of Standards and Technology (NIST) on behalf of Energy Independence and Security Act (EISA) started in 2007 the coordination of a framework that covers communication protocols and standards which will enable the interoperability between intelligent devices and control systems. In June 2009, the NIST published first results of that framework after working with more than 1000 stakeholders [158]. This point includes the communications required by the smart grid’s four priority functionalities defined by the Federal Energy Regulation Commission (FERC) [65], which works together with the NIST and the DOE.

The main leadership in EU is performed by European Smartgrid Technology Platform for Electricity Networks of the Future created in 2005 [52]. It is composed by 4 work groups: Network Assets, Network Operations, Demand&Metering and Generation&Storage. Furthermore, the DG TREN (“Directorate-General for Energy and Transport”) has announced the creation of the “Smart Grids Task Force”, whose first results are going to be presented in May 2010. Many companies which belong to different sectors have seen a great business opportunity and they are currently working to make themselves room in the smart grid’s market. Thus, projects such as Google Power Meter [74] and the ones carried out by companies like Cisco, IBM or Microsoft cannot be forgotten.

The current grid has become outdated and has commonly known problems of inefficiencies and low robustness due to the lack of automation [167] but it could be improved by coordinating processes between Intelligent Electronic Devices (IEDs). Future smart grids must be understood as complex networks of IEDs, wired and wireless sensors, smart meters, distributed generators and dispersed loads that require cooperation and coordination in order to play its expected role. It is needless to say that ICTs, heterogeneous communication networks, trust management and technological integration should play an essential role in this scenario.

The change towards the so called “Intelligent Power Grid” or “Smart Grid” promises to be a change in the whole business model involving utilities, regulation entities, service providers, technology suppliers and electricity consumers. In fact, this transformation towards an intelligent network is possible by importing the philosophy, concepts and technologies from the Internet context. According to the definition in the Strategic Deployment Document (SDD) [52], “a smart grid is an electricity network that can intelligently integrate the actions of all users connected to it (generators, consumers and those that do both) in order to efficiently deliver sustainable, economic and secure electricity supplies”, while at the same time enhancing the integration of the unpredictable supply from renewable energies [44][50].

Generally speaking, a smart grid has some communication requirements according to the characteristics wanted [201][33]:

- **Intelligent**: Able to detect overloads in the system and readdress the energy to prevent or minimize outages. In general, it will be able to deal with problems that require the lowest response time. A detection of emerging problems on the system before they affect the service is a desirable property.
- **Efficient**: Capable of supporting increased consumer demand without adding any extra infrastructure.
- **Accommodating**: Able to accept energy from any energy source including solar and wind energies. It has to be capable of integrating any technology (for example energy storage technologies) when they would be market-proven. It is expected a reduction of the Operating Expenses (OPEX) by lowering capital investment in generation, transmission and distribution assets by smoothing out power consumption.
- **Motivating**: Enabling real-time communication between the consumer and the utility. This way, consumers can tailor their energy consumption based on individual preferences, like price and/or environmental concerns.
- **Opportunistic**: Creating new opportunities and markets by means of its ability to capitalize on plug-and-play innovation.
- **Quality-focused**: Capable of delivering the necessary power quality (free of sags, spikes, disturbances and interruptions) to empower the increasing digital economy and the data centers, computers and electronics.
- **Resilient**: Increasingly resistant to attacks and natural disasters as it becomes more decentralized and reinforced with smart grid security protocols that protect the grid by incorporating protective systems to secure it against threats. the automation of the grid healing can be obtained by incorporating rapid communications, advanced diagnostics and feedback control which returns the system to a stable state quickly.
- **Green**: Slowing the advance of global climate change and offering a genuine path towards significant environmental improvement. It allows to decrease the peak demand by measuring, controlling and shifting the consumption to lower peak usage and provision of better and quantifiable information to customers.

Smart Grids are being developed at an ever-increasing pace and many entities are focusing on them because of their strategic features, mainly from IntelliGrid [44]. Some of the tasks of Smart Grids are grid automation, distributed energy resources coordination, voltage/reactive power control, billing, contingency analysis, protection management and service restoration, among others. Moreover, one of the main challenges of Smart Grids is to redesign the current PLC network architecture for which broadband PLC is considered a key enabling technology by the IEEE. For this reason there is a close liaison between the IEEE WGs P1901 and P2030 (Smart Grid).

Moreover, the Smart Grid architecture is much more complex to operate than the current utility grid scheme because it is based on a decentralized scheme with elements which are logically identified but not geographically located. These new IED are smart electric meters, automated utility substations and new sensors networks. Smart Grids are characterized by a two-way flow of electricity and information, capable of monitoring and responding to changes in anything from power plants to customer preferences to individual appliances. Therefore, the two-way data communications system will provide electric utilities with real-time visibility and control of the electricity used by customers. These networks will basically manage multimedia user data and control data from established IEDs. Since the architecture has to work over a heterogeneous network which may consist of composed wireless and broadband PLC systems, the QoS parameters must be mapped between these technologies in order to obtain seamless end-to-end QoS.

The Smart Grid has also a much more complex infrastructure than the NGN. Therefore the development of standards for the full vision of the Smart Grids is going to be a multi-year effort. The Smart Grid is really a system of systems because it is usually highly fragmented, owned and operated by many companies. Standards play a fundamental role because they provide a common set of protocols that can provide end-to-end communication over physical and link layer underlying heterogeneous technologies, and can therefore provide interoperability. Smart Grids are expected to be built by basically using Internet standards because of their advantages in terms of ubiquity, implementation, interoperability and cost. The National Institute of Standards and Technology (NIST) has released a draft [158] to identify existing standards that could be applied to Smart Grids. The SNMPv3 protocol for network management has been proposed, however in order to meet the requirements of the ITU-T NGN/GSI QoS control architecture, COPS-PR would really have to be supported.

5.5.1. Smart Grids QoS needs

The coming smart grids will manage lots of real-time information through a data network and they will collect information from established IEDs for the purpose of control. This kind of data network is not exempt from the growing need of QoS. Smart grids are expected to face a drastic increase of information demand, communication and various data such as voice, data, image, video and multimedia communications, which will have to be accessed anywhere at any time.

Smart grids need to communicate many different types of devices, with different needs for QoS over different physical media. Availability is also crucial for the correct operation of the network. The elements of a smart grid, the so called Intelligent Electronic Devices (IEDs), can have very different QoS necessities. For example, real-time communications are required in the case of fault detection, service restoration or quality monitoring; periodic communications are used in Automatic Meter Reading systems (AMR) [223]; bulk data transfers are useful to read logs and energy quality information. Smart grids would not be possible without the existence of the IEDs which can play as sensors and/or actuators. There exist many types of IEDs depending on the function carried out. The most important ones are:

- Signal transmission in order to control a remote mechanism.
- Detection of the current circulation in a wire.
- Operational telephony for maintenance and operational activities between electrical substations and the transformation centers.
- Access to corporative applications.
- Video-vigilance.
- Alarm management (temperature, gas, humidity, inundation, etc.)
- Billing and service programming.

- Wave quality measurement.
- Protections management.
- Supervision of the Telecommunications network.
- Meters management (AMR and others).

The IEDs involved in these processes can be situated in different locations due to the pursued decentralized architecture and can use multiple access technologies For example, electrical substation elements are connected to the substation's Ethernet network; sensors can be installed along electrical cables communicated through wireless sensor standards, for example based on IEEE 802.11s. Communications from the control center to energy meters and between substations can be carried out via a high variety of technologies such as PLC, UMTS, GPRS or WiMAX.

Today, ICT is applied at the transmission and subtransmission level [52]. Different standard protocols at various voltage levels and for different kinds of equipments are used. The medium and low voltage communication assets are characterized by economically limited ICT infrastructures. Standardized, open information models and communication services for all data exchange are needed. Related to this point, a key issue in the future is to adequate communication for new services and players. Smart grid requires major communication infrastructure improvements for data exchanges, technical support services, scalability, traffic prioritization (QoS), security and controllability.

Due to these juncture, smart grids will be supported by highly heterogeneous data network with strict constraints of QoS. Therefore, one of the most important needed specifications for the smart grids are those regarding to the necessary communications to support them, based on standards and strict QoS control for the future management of end-to-end QoS for all grid communications. A framework for management of end-to-end QoS for all communications in the grid will be a must in the future. In fact, a suitable communications infrastructure allows increasing the efficiency of the electric system further than what is possible with automation without communication competencies. Hence, we also propose the use of a QoS Broker for Smart Grid based on the ITU-T's NGN/GSI architecture for the high-level management of the Smart Grid data network which includes the acceptance of traffic streams and the QoS management.

Related to data networks, smart grids consist of three different communication networks [7]: Backhaul, Meter and Home Area (HAN). It can be can added the communication with and among IEDs at the distribution level. Their requirements and design ought to be different due to the dimension of each of them or to the different applications supported by them. For example, HAN is referred to automated systems which allow the user to be proactive in the use or the generation of energy. This kind of network is made up of electrical vehicles, meters, thermostats, energy generation systems and many more.

A communication paradigm based on IP protocol has been proposed for the smart grid in [227], since it is the most widely used protocol for communications. On the other hand several promising standards have recently appeared for Smart Grids which base their communications in the IP protocol [89][159]. An appropriate starting point for further standards development would be the harmonization of IEC 61850 standards as they address communications for DER and ADA. Moreover, the debate on whether it should be deployed on IPv4, with IPv6 support, or on native IPv6 continues. Stakeholders should know that it could help drive the adoption of IPv6 massively in corporate and home networks besides the benefits of being able of having 5.000 homes (and meters) in a single subnet. We may think on many reasons for betting for the latter, like the possibility of having 5.000 homes (and meters) in a single subnet. But the most important reason is that it could help drive adoption of IPv6 massively in corporate and home networks. This would turn the Smart Grids into the longly expected killer application for IPv6 [204] along with the future multimedia applications enabled by the NGN architecture.

5.5.2. QoS mappings

The main advantage of the QoS Broker based architecture for QoS control is that the requester node simply needs to specify the parameters for the QoS Broker (priority, bandwidth and delay). In order to ensure the QoS along the path, border interdomain devices have to map the QoS parameters of each flow to suitable ones as each network has its own QoS specification. Internetworking between different network technologies is crucial, both in SENs and in NGNs, to support end-to-end QoS. QoS Brokers are aware of those mappings and they decide whether the network has sufficient resources when a new request is generated and they reconfigure the mappings on demand if it is necessary.

One of the most important industrial research challenges in the mapping field is that the capabilities and parameters change broadly from technology to technology. The QoS mapping between different network technologies is especially important in Smart Grid because of the idiosyncrasy of this technology. One of the reasons why the IEEE P1901 QoS structure is so similar to IEEE 802.11e is to ease the mapping.

QoS mappings between different networks technologies are broadly categorized as vertical and horizontal (Figure 73) [135]. Vertical class mapping refers to the mapping of QoS parameters between two adjacent layers of the same protocol stack (e.g. layer-2 vs. Layer-3). Horizontal class mapping refers to the mapping between two different technologies of the same layer, but different protocol stacks (e.g. layer-2 WiMAX vs. Layer-2 PLC).

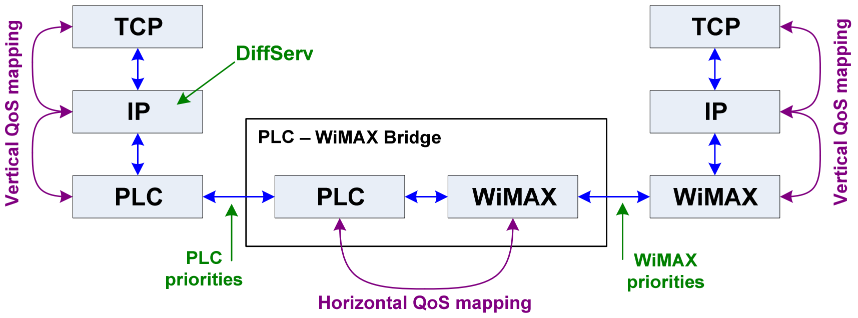


Figure 73: Vertical and horizontal QoS mappings

Horizontal mapping can be carried out by a specialized mapping algorithm between each pair of layer 2 technologies. The advantage of this type of mapping is that accuracy is improved although it does need NxN mappings, where N is the number of different technologies [135]. Horizontal mapping can be also carried out by mapping the QoS at layer 3 which is less precise but only requires a mapping to layer 3 for each layer 2 technology. Even though this latter simplifies the algorithms and would therefore be more suitable for a generic QoS Broker, a definition of the specific mappings between layer-2 Smart Grid technologies would be needed given the intense use of mappings.

The QoS Broker in Smart Grids is aware of the needs of the different data flows of the network regarding priority, bandwidth and delay. The parameters specified by the PIBs of

OPERA PLC and by the PIBs of IEEE P1901 (can also be reused for 802.11e) affect the total delay of the data flows in its segment of the network and there is a balance between bandwidth efficiency and minimizing the delay. The QoS Broker decides how to balance these parameters in order to meet the delay constraints in the most efficient way by taking into account the flows in the network, the network topology and the customers SLAs. When the QoS Broker copes with a long period of loss of quality in a link segment, it ought to reconfigure QoS parameters and it could tear down some accepted low priority flows if needed. It turns out that a problem of both IEEE 802.11 and PLC is the possibility of a sudden loss of quality of the signal in the medium. For the purpose of admission control and parameter recalculation, the QoS Broker must use a model that can relate the parameters of each technology with the parameters bandwidth and delay. This process must be carried out taking into account the knowledge of the network topology as the end-to-end parameters have to be calculated too in order to meet SLAs. For example, the minimum bandwidth of a path can be calculated by checking that all the networks along the path have the required bandwidth and the maximum aggregated delay by summing all the delays of layer-2 segments. The problem of calculating the end-to-end QoS through different protocol stacks and technologies has previously been addressed [34]. The models presented in [9] for IEEE 802.11e and [182] for OPERA PLC are used for the calculation of the delay and bandwidth in each domain.

QoS mappings between OPERA PLC and 802.11e

In the Smart Grids, there is real-time traffic such as the traffics related with protections, restoration processes of service or real-time pricing. Also lower priority traffic flows coexist such as meter reading. A mapping table can be defined to map priorities from Diffserv (DSCP) to OPERA-PLC and IEEE 802.11e priorities as shown in Table 17. Although the mapping between priorities is static, the parameters of each priority are dynamic and controlled by the QoS Broker. In the case of IEEE 802.11e (and IEEE P1901), the Smart Grid QoS Broker controls the parameters of ACs, which are also statically mapped to User Priorities (UPs) as defined in IEEE 802.11e. Other direct mappings could be used if needed, such as the IEEE 802.1p priorities if Ethernet technology is also to have a role. Some examples of use could be inside an electrical substation based on IEC 61850 family standards [89]) or in the aggregation network at the IP backhaul.

Traffic type	DiffServ (DSCP)	PLC priority	802.11e (UP/AC)
Critical	EF	7	7 / AC_BK
Telecontrol	AF41	6	6 / AC_BK
Voice	AF34	5	5 / AC_BE
Video	AF31	3	3 / AC_BE
Meter reading	AF21	1	1 / AC_BI
Bulk	BE	0	0 / AC_BI

Table 17: Example of priority mapping between DSCP, PLC and IEEE 802.11E

The different traffic categories for the network in Table 17 are:

- **Critical**: Refers to critical communications of protections and service restoration.
- **Telecontrol**: Refers to other automation messages.

- **Voice**: Communication between the control center and technicians.
- **Video**: For surveillance purposes.
- **Meter reading**: To read meters from final users for billing purposes in AMR scheme.
- **Bulk**: For reading logs or uploading files to devices.

The mappings between IEEE P1901 PLC and IEEE 802.11e are direct as the QoS facility of the former is based on the QoS framework of the latter.

5.6. Chapter summary

The new generation of broadband PLC technologies is in their final phase of standardization. These technologies have improved to the point of being considered as NGAs. When considering the QoS requirements presented in the introductory chapter, the need for an architecture to manage seamless end-to-end QoS through broadband PLC networks which are able to deal with multimedia data traffic remains clear.

This chapter has reviewed the OPERA project specification and the IEEE P1901 draft standard proposal for broadband PLC technologies, and it has shown how the end-to-end QoS resource control architecture developed by the ITU-T's NGN/GSI can be applied to broadband PLC and Smart Grid communication network requirements in order to provide QoS management in a centralized and standardized manner. Furthermore, this chapter has demonstrated how the QoS Broker can also be used to handle layer-2 QoS, besides the layer-3 QoS, and to improve the efficiency of communications.

The QoS Broker uses COPS-PR for intra-domain policy delivery and if it intended to use the ITU-T's NGN-GSI architecture in broadband PLC networks, a PIB for each one of the technologies under management would have be defined. A PIB for the OPERA proposal and another PIB for the IEEE P1901 proposal have been proposed and the hierarchy of parameters that the QoS Broker can manage using the COPS-PR protocol has been discussed.

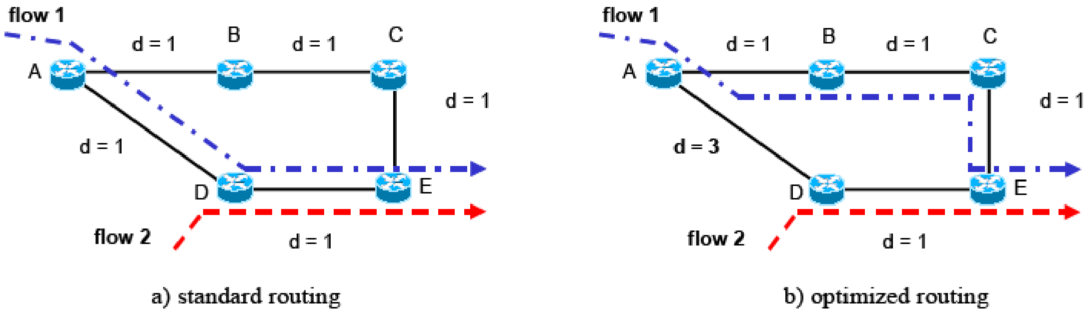
Finally the increased need for Quality of Service of the coming Smart Grids communication networks has been highlighted and a description of how the QoS Broker architecture can be applied to cope with those more stringent QoS requirements has been provided. The importance of QoS mapping between different network technologies along the heterogeneous path of end-to-end multimedia communications and inside the Smart Grid has been underlined and a method for priority class mapping from OPERA PLC to IEEE 802.11e has been proposed.

These contributions to the thesis have been presented in [211], [224], [226] and [227].

6. TRAFFIC ENGINEERING WITH ARTIFICIAL INTELLIGENCE IN NGNS

6.1. Introduction

Traffic engineering enables operators to reduce the global cost allows reducing the operators' global cost of operations by optimizing the available bandwidth. This technique prevents situations where some parts of the operator network are congested while some others remain underused, mainly due to fluctuating traffic demands which can occur even in well dimensioned networks [217]. Routing optimization is a traffic engineering method that provides a possible solution to this problem. Routing optimization is a key aspect to take into account when providing QoS in next generation networks (NGN), especially in access networks.



a) standard routing

b) optimized routing

Figure 74: Example of routing optimization

The aim of routing optimization is the optimization of networks so that more traffic can be routed. One way of achieving it is to modify the link weights and therefore the metrics. The problem of weight setting with conventional link state routing protocols for routing optimization has been studied in order to adjust link utilization and it has been object of study of a few authors. Depending on the dimensions of the topology, this weight setting problem may become a NP-hard problem [67][70], which can be solved through heuristics with artificial intelligence techniques. Among different approaches, GAs have been devised as one of the most appealing methodologies to tackle this problem since they become NP-hard when applied to large networks. In particular, some authors have used hybrid GAs (memetic GAs) which incorporate local search procedures in order to optimize the GA results.

In order to optimize routing of a NGN with these requirements, one option is to use an architecture similar to the one shown in Figure 75, where a decision server applies an offline routing protocol over the known network topology and uses an artificial intelligence algorithm to decide the optimal weight setting of the links which are later delivered to the nodes. The remarkable parallelisms with the centralized QoS management based on policies (PBNM) are particularly relevant. Therefore, in this chapter an offline routing optimization will be implemented in the QoS broker in order to apply traffic engineering to the OSPF (Open Shortest Path First) routing of the IPv6 testbed.

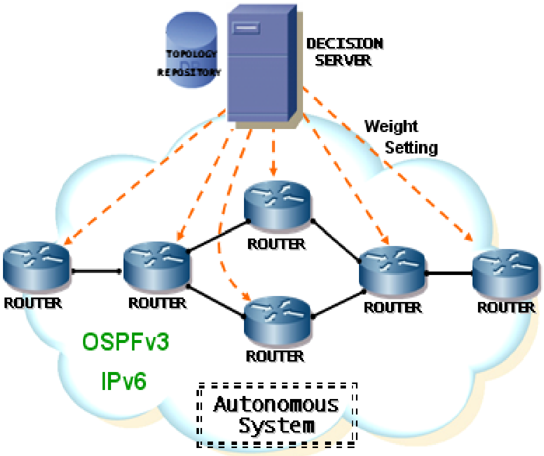


Figure 75: Network architecture for an offline routing optimization in NGNs

6.1.1. Objectives

The purpose of this line of research is to develop a system to optimize the routing of NGNs with HGA algorithms. This system has to successfully be applied to the IPv6 testbed. In order to achieve these objectives, the first step is to perform a comparative analysis of the main hybrid genetic algorithms (HGA) proposals as well as comparing them with other algorithms for the same problem by means of simulations. To do so, a previous study of the state of the art must be done.

The second objective is to enhance the QoS broker by developing an application to apply the HGA algorithm, chosen from the results of the analysis, to the off-line OSPFv3 routing protocol, thereby achieving new genetically optimized link weights. This application must also deliver the new weights to the IPv6 testbed.

The last objective is to validate the correct performance of the system by modifying the testbed in order for the nodes to accept new weights from an external agent. The first performance test will be carried out over a testbed with commercial Cisco System routers running OSPFv3 routing protocol in order to validate the communication with the QoS broker which will deliver the new weights with SSH. The second test will be carried out with the GNU/Linux testbed using the COPS-PR protocol for the weights delivery, according to ITU-T specifications for the NGN-GSI architecture.

6.2. Traffic engineering

Traffic Engineering [9] is a set of principles, architectures, and methodologies for resource utilization optimization rather than a QoS architecture. Without TE, traffic may concentrates into the same nodes at the same time. Therefore, network may have some congestion points while the load in the whole network is not very high. This type of congestion decreases the performance criteria of the whole network while the total throughput is low.

TE distributes traffic in an optimum manner in order to minimize the congestion point in the network, thereby improving the performance of the whole network in terms of minimizing traffic loss, minimizing delay, maximizing throughput, minimize maximum congestion, and enforcing SLAs. A classical block diagram of traffic engineering system can be viewed in Figure 76.

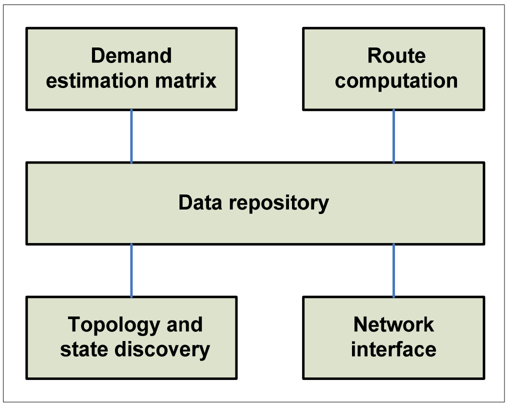


Figure 76: A traffic engineering system

The topology and state discovery block monitors any change in network topology and link state. Dynamic information such as residual bandwidth or link utilization may be collected. To distribute topology and resource availability information, one approach is using OSPF routing protocol. The demand estimation block estimates the demand of user in term of traffic, thereby estimating the traffic load of the network. This estimation usually relies on SLAs but it can also rely on traffic measurements. The route computation block calculates route based on traffic demands. This is the central module of traffic engineering to obtain the objective of resource optimization or policy constraints. Constraint-based routing (e.g. QoS routing, route optimization algorithms) is the main tool for TE to compute routes. Once the route computation block has worked out the optimal routes for the demands, the network interface module is responsible to configure the network elements in the network accordingly. MPLS can be used to establish computed explicit routes. The data repository provides a central database that the mentioned-above modules can store, access and exchange information.

As TE can calculate routes which match specified QoS criteria, it can be used to provide QoS. For instance, QoS requirements specified in SLAs such as bandwidth, delay, or jitter are provisioned in appropriate routes calculated by Traffic Engineering.

6.2.1. Routing optimization fundamentals

In traffic engineering, a network optimal performance is generally accepted as being one where network congestion has been minimized in all the links so that all of them are equally congested. This is done from the existing resources utilization in a domain and the traffic demand matrices. One possible way of optimizing this performance is to manipulate the routing process of the packets, that is to say, by modifying the routing protocol.

The function of the routing protocols is to find the best route between an origin node and a destination node, from a minimization of the cost function. Given that the metrics provide a comparative measure to decide which path is better, we must make sure that the values of their components are properly adjusted so that they improve the global performance. Therefore to optimize routing we need to define an objective function which takes into account the link's usage in a quantifiable way and hence considers the routing cost.

The general routing problem can be described as the problem of optimizing the minimization of an objective function from a given network topology and a traffic demand matrix. Under stationary conditions the problem can be solved with linear programming (LP) and the result will be the best possible routing for all the possible flows so all the traffic will be globally optimized for all the networks in the domain.

Existing research lines tend to use the OSPF and ISIS protocols. The advantage of these protocols is that they incorporate IPv6 versions broadly used by commercials routers (OSPF for IPv6 in RFC-5340 [29] and IS-IS for IPv6 in RFC-5308 [85]). Moreover both protocols have versions oriented to traffic engineering with support for IPv6 (TE extensions to OSPFv3 in RFC-5329 [108] and TE extensions to IS-IS in RFC-5305 [30]). These characteristics make them the most commonly-used protocols in NGN-related research. In this thesis it has been used OSPF, in concrete its version 3 for IPv6 [29].

In order to fully understand how an effective approach to optimize overall network performance with OSPF and before considering the options available to achieve this optimization, it must first explained the how the OSPF protocol behaves.

6.2.2. Link state routing

The OSPF protocol is one of the most common link state routing protocols in packet networks. This protocol works with a metric based in a cost value associated to each link and applies Dijkstra's shortest path algorithm [39] to find out the shortest path to each network based on these costs. In OSPF the metric of a path to a given network is the sum of all the costs to that network.

In the case where multiple paths exist to destination with equal metrics, OSPF can balance the load with equal cost multiple path (ECMP) so that the traffic flows will be theoretically evenly split between all the paths with the same metrics. Even though it is typically impossible to guarantee an exact even split of the load, it was decided to compare OSPF with ECMP and OSPF without ECMP using an exactly even traffic distribution to achieve the simulations of thesis.

Once the initial convergence phase has finished, any change which occurs in any link, for example a weight modification, will result in only the affected link's modification being flooded. Each one of the routers will have to decide whether or not to recalculate all the information in the routing table. Therefore if the number of changes of the link weights is high, it may lead to an inefficient use of the net's resources, as well as in bandwidth as in CPU.

In this thesis the inverse of the link capacity has been used as a default configuration of the link weights (or costs). This methodology was first proposed by Cisco [27] and according to [67][68] is the best way to adjust the link weights in default configurations with OSPF.

6.2.3. The OSPF weight setting problem

Given a known network topology and a predictable traffic demand, the OSPF weight setting problem (OSPFWS) is to find a set of weights which optimize the network performance and therefore minimize the cost function [67][70]. This problem, as stated previously, can be NP-hard depending on the dimensions of the topology. As this kind of problems can not be solved in a polynomial time, then heuristic search methods must be employed to find the most optimal solutions.

The use of local search heuristics, which apply an iterative process to solve the problem, only work for medium sized networks at the most and they do not guarantee the best possible solution [185]. The most commonly-used algorithm is the one proposed by Fortz and Thorup [68], which proposes minimizing a function that summarizes all the link weights so that it optimizes the global performance of the domain. This proposed cost function is convex, incremental, lineal, continuous and piece-wise, which assigns low costs to the infrequently used links and high costs to the overloaded links. If the problem needs to be solved for medium-big sized networks it is necessary to use artificial intelligence, given that using a local search heuristic is not viable in computing time terms.

6.3. Genetic Algorithm heuristics for routing optimization

Genetic algorithms [83][72] are methods for search, optimization and machine learning which are inspired by natural principles and biology. Differently from other optimization methods, GAs do not assume any structure or underlying distribution of the objective function and employ random, local operators to evolve a population of potential solutions. Since GAs have demonstrated to be able to solve complex problems that previously eluded solution [73], it has been chosen to adopt this optimization model in our design. In the following, it is first presented the basic mechanics of GAs and then explained different GA implementation to solve the routing optimization problem.

6.3.1. Mechanics of Genetic Algorithms

GAs evolve a population of individuals, where each of them represents a potential solution to the problem. Analogous to genetics, individuals are represented by chromosomes, which encode the decision variables of the optimization problem with a finite-length string. Each of the atomic parts of the chromosome is referred to as genes, and the values that the gene can take are addressed as alleles. To implement the principles of natural selection and competition among candidate solutions, GAs incorporate an evaluation function that gives a certain value of fitness to each individual, which indicates the quality of the given individual.

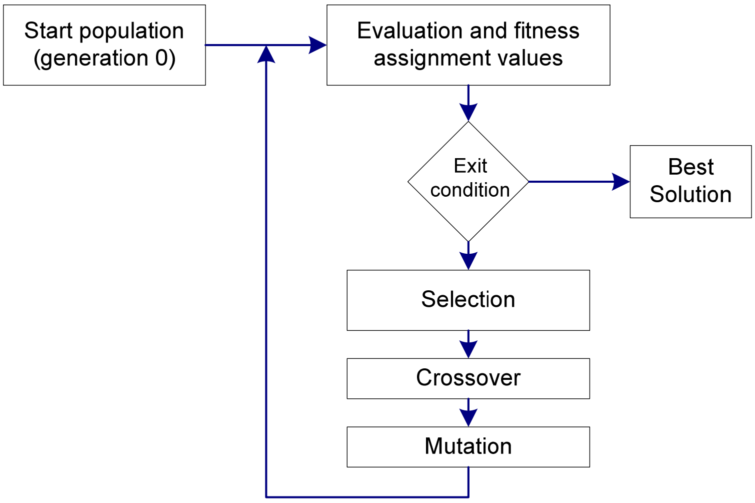


Figure 77: Flow chart of a Genetic Algorithm

Then, this population of individuals, which is usually initialized randomly, is evolved by a continuous process of selection, crossover, mutation, and replacement of individuals. That is, firstly the selection operator chooses the fittest individuals in the population, simulating the survival-of-the-fittest mechanism. Then, the crossover operator takes two or more of the selected individuals and recombines their genetic information in order to generate new, possibly better offspring. Afterwards, mutation introduces random errors on the transference of genetic information from parents to children, and finally, the offspring population replaces the original one. This process is repeated until a stop criterion is met; usually, the process is run during a prefixed number of iterations. Figure 77 schematically illustrates this process.

The synergy of all these operators pressures toward the evolution and selection of the best solutions, which are recombined yielding new promising offspring. In [73], Goldberg emphasized the idea that, while selection, crossover, and mutation can be shown to be ineffective when applied individually, they might produce a useful result when working together. This was explained with the fundamental intuition of GAs, which supports the following two hypotheses. The first hypothesis is that the combination of the selection and crossover operators introduces a process of innovation or cross-fertilizing by generating new solutions from the fittest individuals in the population. As a consequence, new individuals are expected to be different from and better adapted than their parents. The second hypothesis is that the combination of selection and mutation represents a process of continuous improvement or local search. Thence, this process searches around the best solutions in the population with the aim of finding better solutions that are close to the parents.

6.3.2. Design of a Genetic Algorithm for routing optimization

With the basic mechanisms of GAs in mind, now we are in position to proceed with the description of how GAs have been applied to the routing optimization problem. In what follows, we present the typical representation employed by several authors, and discuss which types of genetic operators have been used in different approaches [17][47][143][146][175].

In order to solve the OSPFWS problem, the solution must contain the weights of each link of the network. Therefore, the individual is typically represented as a vector that contains all the link weights of the domain, which range in [1, wmax]. Thus, each individual provides a complete solution to the problem. This population is typically initialized with the default weight setting with the inverse capacity procedure as proposed by Cisco [27]. On the other hand, several evaluation functions that provide a measure of the link performance of the domain, showing the most overloaded or the global average performance, have been used by different authors. In the following section three of these functions are explained in more detail.

Different genetic operators have been used in different GA implementations for this problem so far. Two selection schemes were employed in the approaches studied in this thesis: rank selection and proportionate or roulette wheel selection. Rank selection ranks the individuals of the population according to their fitness, and those with better ranking are selected to be in the next generation. An especial case of rank selection is tournament selection, which uses a set of s randomly selected individuals for ranking instead of considering the whole population. On the other hand, proportionate selection gives each individual a selection probability that is proportional to its fitness with respect to the fitness of the other individuals in the population. Optionally in this phase a technique named elitism can be used, which consists in always passing at least one copy of the best chromosome to the next population, so the best individuals are not lost due to the effect of the genetic operators.

Then, crossover and mutation reproduce the parent population as follows. Crossover is applied to each pair of parents with a certain probability. If applied, crossover randomly generates a cut point and uses this cut point to shuffle the genetic information of the parents. Therefore, crossover creates two new individuals that mix the genetic information of the parents. If crossover is not applied, the offspring are exact copies of the parent. Thereafter, the offspring undergo mutation. That is to say, for each gene, a random number is generated and, if it is lower than the probability of applying mutation, the gene is mutated by assigning a new randomly selected value in the interval [1, wmax]. After mutation, the new population replaces the original one.

A representative example of a GA applied to routing optimization is provided by Ericsson et al. [47]. In this case the GA is based on the idea proposed by the heuristic search in [67][70] and it applies the same cost function. The representation of the population’s individuals is formed by the set of all the link weights in a vector. The population initialization is randomly generated and the selection method is the rank selection where it divides the population in three sets of α=20% (elitism), β=70% (crossover) and γ=10% (discarding), in respect the total population size between 50 and 500 individuals. The crossover probability is 70%, mutation probability is 1% and the number of iterations is variable between 500 and 700.

Throughout this section, it has been explained the process organization of GAs, have intuitively discussed how and why they work, and have shown how GAs have been applied to the routing optimization problem. In the next section, we take these ideas and explain how GAs can be enhanced by incorporating a new local search procedure that enhances the original local search mechanism of GAs – that is, mutation – in order to converge quicker to the objective.

6.4. Local search with GA for routing optimization

The hybrid genetic algorithms (HGA) [143], or memetic algorithms, are distinguished from the GA because they append a local search heuristic applied during the evolutionary cycle, as can be seen in a typical HGA flowchart in Figure 78. The objective of this local search procedure is to improve the effectiveness and efficiency of a GA when converging to an optimal solution of the problem. In this section we provide an inedited theoretical comparative analysis of the main three HGA proposals.

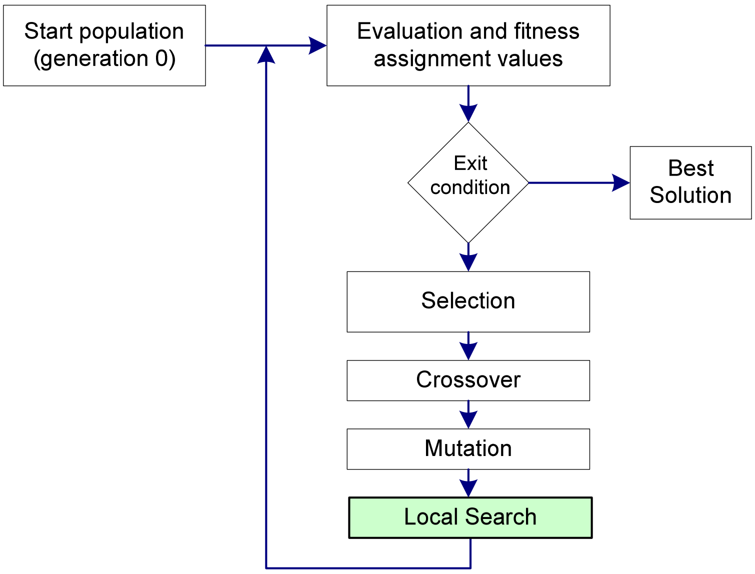

Start population (generation 0)
Evaluation and fitness assignment values
Exit condition
Best Solution
Selection
Crossover
Mutation
Local Search

Figure 78: Flow chart of a Hybrid Genetic Algorithm

The hybrid genetic algorithms obtain better results when optimizing the global performance of the domain with OSPF routing process rather than the simple GA [17][146][175]. Some benefits that the local search addition provides are acceleration in the optimization process (in computational time) and improvement in the quality of the solutions, which are more optimized, that is to say better, as demonstrated in [17] and [174] (Figure 79). However, a disadvantage is the potential loss of the global maximum, getting stuck in a local maximum, as happens when there is an abuse of the genetic operators.

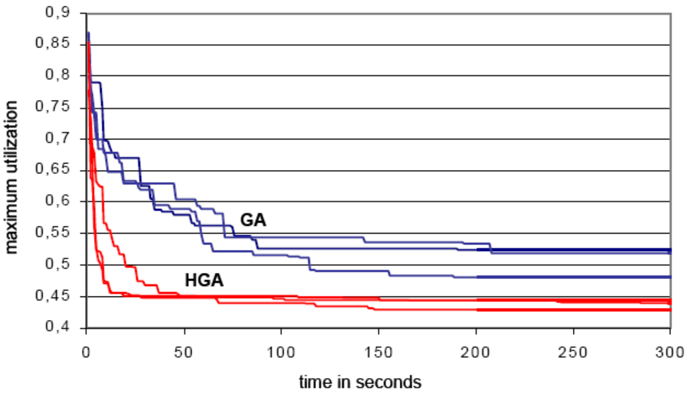


Figure 79: Comparison of HGA and GA [174]

A theoretical analysis of the HGA algorithms proposed in [17], [146] and [175] has been carried out. Table 18 presents a summary of the most representative genetic parameters of the three proposals. All of them have used the same values in the crossover parameter and mutation, but they differ in the selection method, weight representation and overcoat, where to apply the local search procedure and the fitness function.

	Mulyana & Killat [146]	Buriol et al. [17]	Riedl & Schupke [175]
Population Size	50	50	20
Chromosome Representation	Set of all domain's link weights in a vector		
Weight Representation	[1, 99]	[1, 20]	[1, 20]
Number of Iterations	200	200	200
Selection Method	Rank Selection: α=20%, β=70%, γ=10%	Rank Selection: α=25%, β=70%, γ=5%	Roulette Wheel Selection
Crossover Probability (P_c)	0.7	0.7	0.7
Mutation Probability (P_m)	0.01	0.01	0.01
Heuristic Search	Best individual	Individual generated from the crossover	All individual

Table 18: Parameters of the Hybrid Genetic Algorithms

The fitness function used by Buriol et al. in [17] is the same as the one used by Resende et al. in [172], in the first application of a GA to the optimization problem and has been introduced in the previous section. Both of them use the convex cost function proposed by Fortz and Thorup [68].

$$\min \Phi = \sum_{a \in A} \Phi_a(l_a) \tag{1}$$

subject to:

$$\sum_{u:(u,v)\in A} f_{(u,v)}^{(s,t)} - \sum_{u:(u,v)\in A} f_{(v,u)}^{(s,t)} = \begin{cases} -d_{st} & if\ v=s, \\ d_{st} & if\ v=t, \\ 0 & otherwise, \end{cases} \qquad v,s,t \in N \tag{2}$$

$$l_a = \sum_{(s,t)\in N\times N} f_a^{(st)} \qquad a \in A \tag{3}$$

$$\begin{aligned} \Phi_a(l_a) &\geq l_a & a \in A \\ \Phi_a(l_a) &\geq 3l_a - 2/3c_a & a \in A \\ \Phi_a(l_a) &\geq 10l_a - 16/3c_a & a \in A \\ \Phi_a(l_a) &\geq 70l_a - 178/3c_a & a \in A \\ \Phi_a(l_a) &\geq 500l_a - 1468/3c_a & a \in A \\ \Phi_a(l_a) &\geq 5000l_a - 16318/3c_a & a \in A \end{aligned} \quad (4)$$

$$f_a^{(s,t)} \geq 0, \qquad a \in A; s,t \in N \quad (5)$$

Later, Mulyana and Killat [146] tackled the General Routing Problem by minimizing the following fitness function, which takes into account a weighted addition of the global average and the maximum link utilization.

$$\min\left\{(a_t \cdot t) + \frac{1}{|E|}\sum_{ij}\sum_{uv}\frac{l_{ij}^{uv}}{c_{ij}}\right\} \qquad \forall(i,j)\in A, \forall(u,v)\in VxV \quad (6)$$

$$\delta_{un} f_{uv} + \sum_{m\in V} l_{mn}^{uv} = \delta_{nv} f_{uv} + \sum_{m\in V} l_{nm}^{uv} \qquad \forall(u,v)\in VxV, \forall n,m \in V \quad (7)$$

$$\sum_{uv}\frac{l_{ij}^{uv}}{c_{ij}} \leq t \qquad , l_{ij}^{uv} \geq 0, \forall(i,j)\in E, \forall(u,v)\in VxV \quad (8)$$

Finally, the fitness function employed by Riedl and Schupke in [175], only considers minimizing link maximum utilization with the application of an exponential scalability factor. This way, routing solutions with smaller maximum link utilization receive higher fitness values and a greater chance to be reproduced in the new generation.

$$fitness = \left(\frac{1}{\rho_{\max}}\right)^p \qquad , p > 0 \quad (9)$$

subject to:

$$\rho_{\max} \geq \rho_{ij} \qquad \forall(i,j)\in A \quad (10)$$

$$\rho_{ij} = \sum_{u\in V}\frac{f_{ij,u}}{c_{ij}} \qquad \forall(i,j)\in A \quad (11)$$

6.5. Evaluation of the HGAs

The aim of this section is to choose one of the HGAs proposed in the previous section and consequently implement it in a real testbed, providing a comparative analysis of the main three HGA proposals. The main problem when comparing the different algorithm's proposals of the authors is the fact that each one uses its own network topologies and traffic matrices. As the objective is to evaluate comparing the algorithms it is necessary to determine a common topology and traffic demand matrix to apply and test them. When this comparative was first proposed by the authors in [207] there was a physical limitation of twelve routers creating a restriction regarding the topology of the test. Therefore the network topology N11 used in [175] has been selected. This network topology, which has 11 nodes and 48 unidirectional links, can be seen in Figure 80.

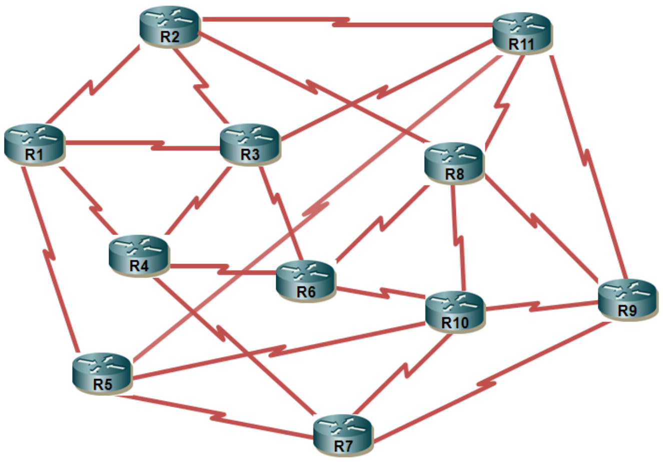


Figure 80: Network topology N11 used to evaluate the routing optimization

In order to evaluate the proposals comparisons between the different approaches have been made: the default inverse capacity metric (InvCap) [27], the local search algorithm proposed by [70], the GA proposed by [47] and the three HGAs proposed by [146], [17] and [175]. The parameters proposed by the authors of the HGA algorithms have been used in this evaluation. Algorithms with a maximum number of iterations of 250 and a maximum percentage of total weight change of 30% from initial configuration were used in order to avoid excessive flooding which would have consequently led to excessive routing re-calculations. Six traffic demand matrices with increasing traffic have been generated.

The graphs show how, in general, the HGA algorithms improve local search and simple GA algorithms' results. The two graphs, OSPF without ECMP (Figure 81) and OSPF with ECMP (Figure 82), show how the HGA algorithm proposed by Buriol et al. is the one that most effectively minimizes the maximum usage of the most congested link. Overcoat in the worst traffic cases when under default conditions (InvCap) there is link overloaded, even though there exist minimum differences with respect to the other two HGA proposals.

On the other hand the HGA algorithm proposed by Mulyana and Killat is the one which best minimizes, albeit marginally, the average usage of the links in all cases, as can be seen in the graphs of OSPF without ECMP (Figure 83) and OSPF with ECMP (Figure 84). In the case of average usage of the links, even though the HGA algorithms always provide better results, the deviation with the others are minimal in percentage.

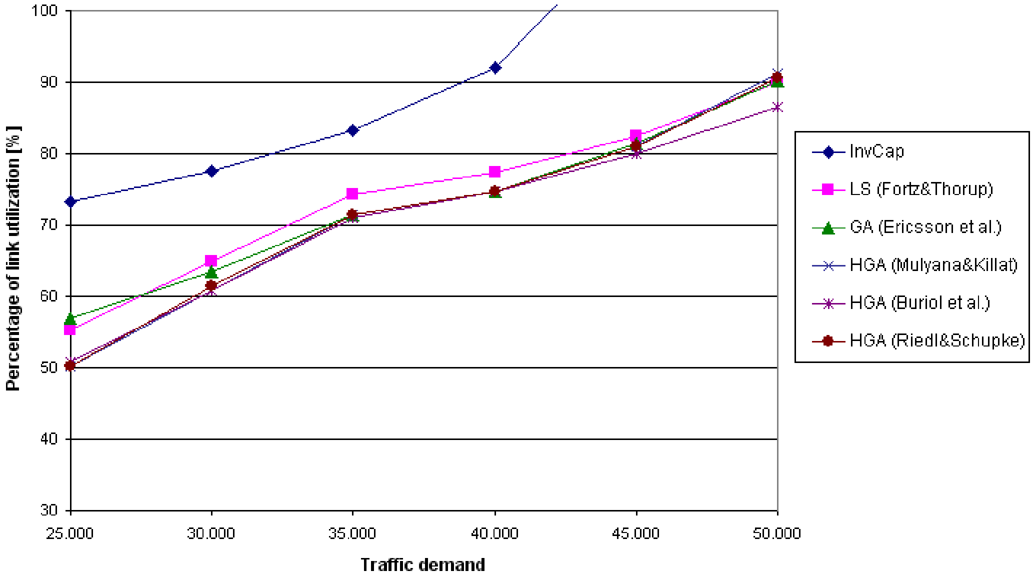


Figure 81: Maximum link utilization without ECMP

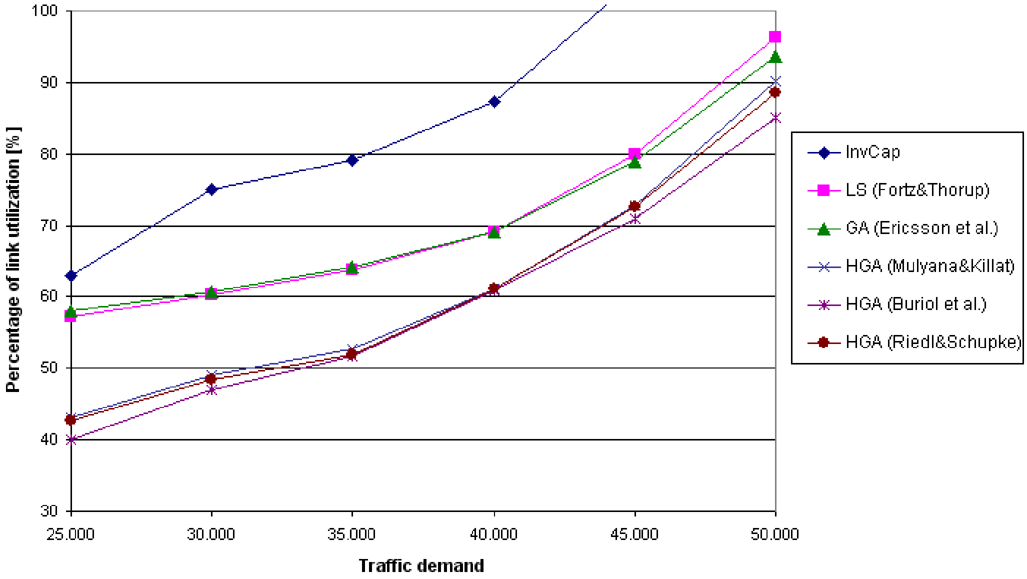


Figure 82: Maximum link utilization with ECMP

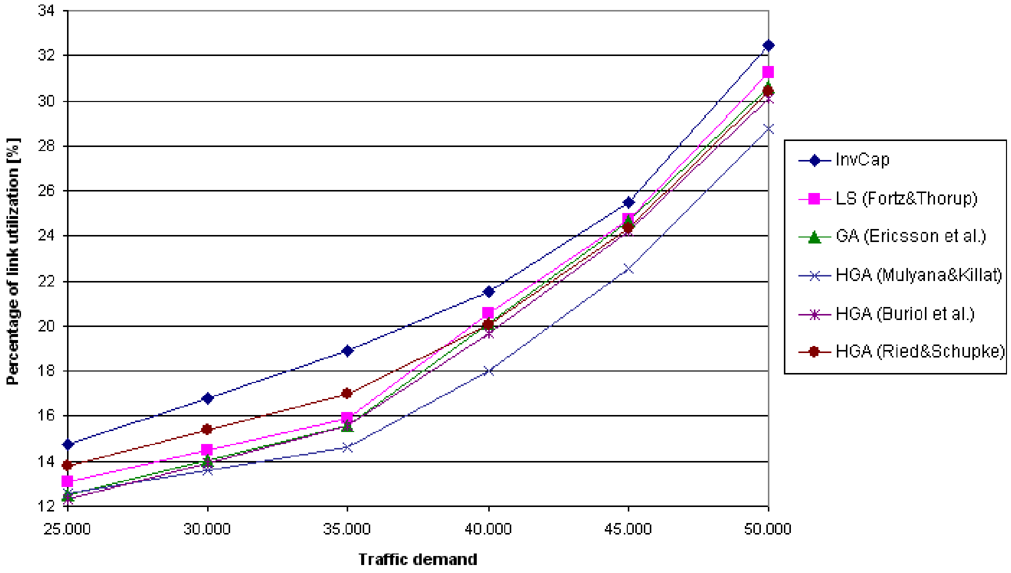


Figure 83: Average link utilization without ECMP

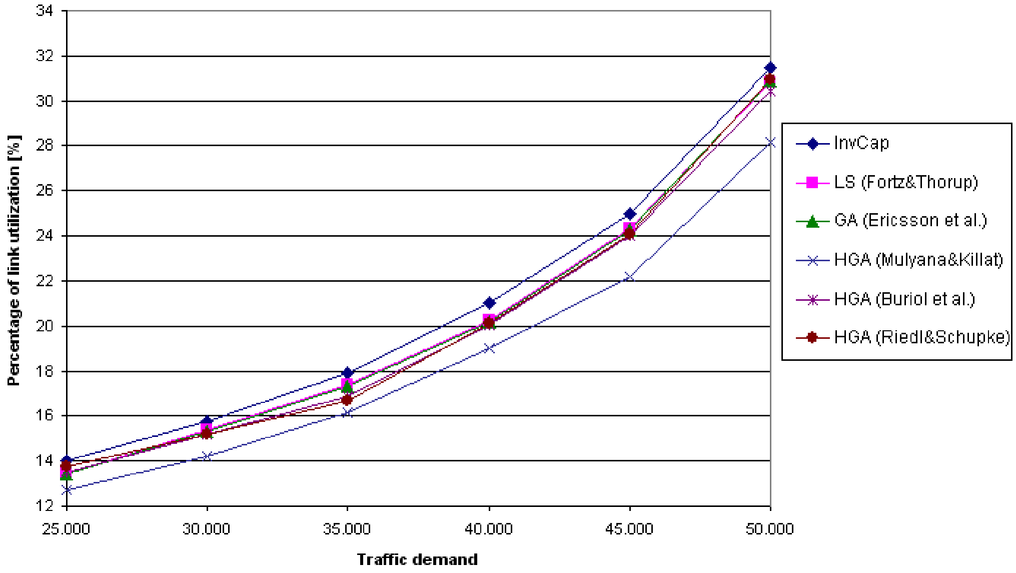


Figure 84: Average link utilization with ECMP

According to results minimal variations between the three HGA proposals exist, both in the average usage as well as in the maximum usage. It was decided to use the algorithm proposed by Buriol et al. because it presents slightly better results in this latter aspect.

6.6. Testbed implementations

Different testbeds have been used to implement the proposed HGA algorithm. The first tests have been done in a testbed with commercial Cisco Routers in order to facilitate the testing. These first tests have been done initially over a 7 node testbed (Appendix D) to easily test the QoS broker applications to send the new weights via SSH instead of sending them to the simulator. Afterwards, N11 network topology used in the simulations has been tested. The final tests have been done over the GNU/Linux IPv6 testbed.

6.6.1. Traffic optimization using Cisco Systems routers

A real testbed with eleven routers has been implemented. Five of the routers are Cisco routers 2621 (R1-R5), with IOS “c2600-is-mz.123-22.bin” and 64M of RAM memory and 32M of flash memory and six routers 2621XM (R6–R11), with IOS “c2600-advipservicesk9-mz.123-22.bin” with 128M of RAM memory and 32M of flash memory (Figure 85). All eleven routers have two WAN Interface Cards 2A/S and two FastEthernet interfaces each. In order to use the OSPFv3 routing protocol, “IPv6 Routing: OSPF for IPv6 (OSPFv3)” of the Cisco’s IOS [28] was used, which is based on the RFC-5340 [29]. The UDP traffic sources used to carry out the tests have been generated with Iperf 2.0.2.

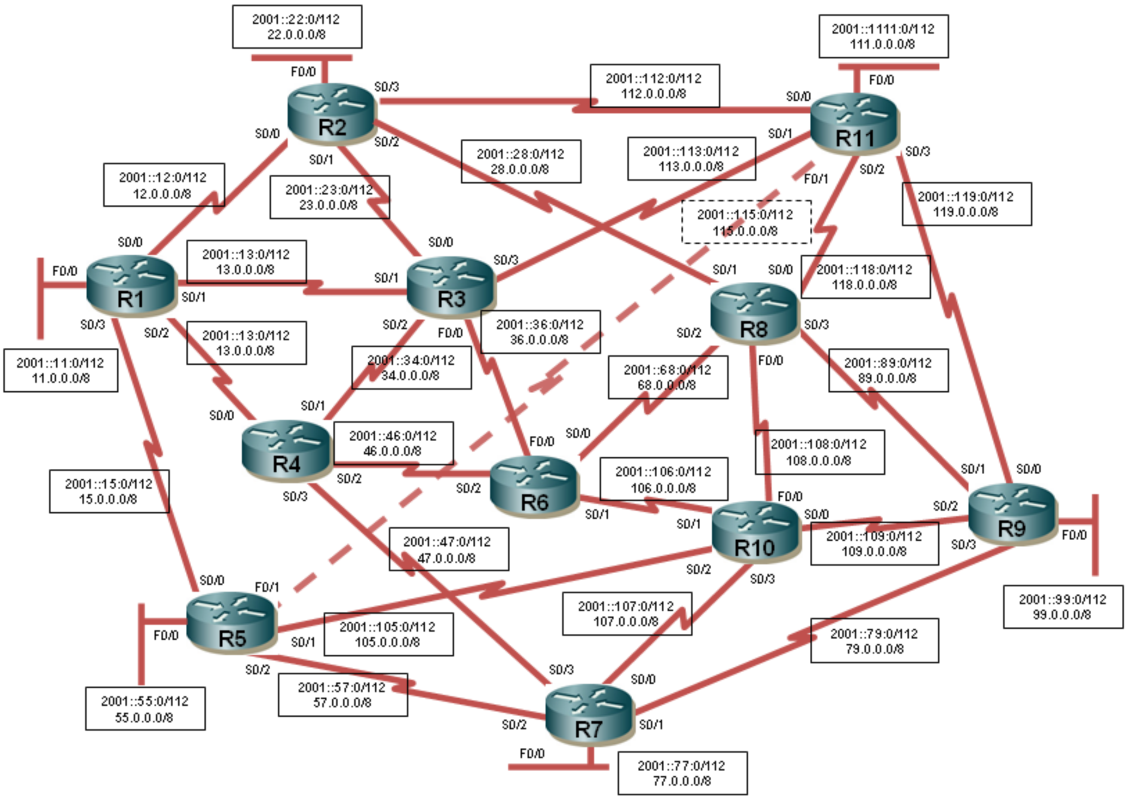


Figure 85: Testbed implemented to evaluate the routing optimization using Cisco Sistems routers

The centralized management of the architecture was created with the previously implemented QoS broker device. This QoS broker has a PostgreSQL database with the domain topology (nodes, link’s capacities and costs) and a client’s traffic demand matrix. Therefore it is a maximum traffic configuration in order to assure client’s contract traffic.

An application has been developed in the QoS broker to compute the offline routing from the domain’s topology and the traffic matrix with the OSPFv3 protocol. The default link weights of the Cisco interfaces are obtained by using the link’s inverse capacity [27] and applying the HGA algorithm proposed by Buriol et al. these link weights can be modified, to obtain the new “genetically” optimized weights. The sending of these new OSPF weights to the Cisco router’s interfaces was done via SSH.

The load balancing in Cisco routers is automatically activated if the routing table has multiple paths to a destination. In this testbed per-destination load balancing has been used, where the router distributes the packets based on the destination address. Another option could have been per-packet load balancing which guarantees equal load across all links but there is the possibility that the packets may arrive out of order at the destination if differential delay exists within the network and it is a processor intensive task which may impact the overall forwarding performance. We should underline the fact that even though per-destination load balancing is an improvement over per-packet, it is still not very good because if substantially more packets are sent to one destination than to another, the overall bandwidth utilization will be uneven.

The results show how initially, as shown in the simulation with the inverse capacity weights (InvCap), the maximum overload was in the R2-R8 link which gets the first to 100% (255/255 txload in the S0/1 interface of R8. When the HGA is applied with the activation of the new module, this one sends the new weights computed offline to the routers. Once the recomputation has taken place, the R2-R8 link lowers to 46% (115/255 txload in the S0/1 interface of R8) and the new most used link is now the R3-R4 link which gets to 51.76% (132/255 txload in the S0/1 interface of R4).

Given that the main function of the QoS broker is to manage QoS DiffServ model policies, in [69] a description of how these kind of optimizations can be applied to the DiffServ model is provide. This further testing has not been undertaken.

In this section a successful implementation of routing optimization in an IPv6 domain with eleven commercial 262x Cisco routers and with a centralized GNU/Linux device. The task of this centralized device was computing the link weights through a HGA algorithm and sending them to domain routers via SSH, where they were configured automatically (Figure 75). The routing protocol used was OSPFv3.

6.6.2. Traffic optimization using GNU/Linux routers

This section presents a real testbed implementation in which the routing optimization presented in the previous section has been upgraded to meet the ITU-T requirements. Thus, in order to deliver the link weights to the routers over a NGN testbed, the COPS-PR protocol has been used.

The application developed in this chapter has been integrated into the ITU's NGN-GSI testbed presented in chapter 4. In this latter chapter the COPS-PR protocol was used to manage the intra-domain policies and resources with the GNU/Linux routers (Rw, Rc and Rn interfaces). These devices are formally the RACF entity and have been named as QoS Brokers (QoSBv6).

The application for computing the offline routing from the domain's topology and the traffic matrix with the OSPFv3 protocol was already installed in the QoSBv6 as seen in the previous section. The default link weights of the GNU/Linux router interfaces are obtained by using the link's inverse capacity [27] and by applying the HGA algorithm proposed by Buriol et al. [17], these link weights can be modified to obtain the new "genetically" optimized weights. The sending of these new OSPF weights to the GNU/Linux router interfaces is now carried out via the COPS-PR protocol.

This way, whenever the administrator decides to optimize routing, which is relatively infrequent, the application applies the HGA to the current topology and sends the new weights to the system Database, which will be copied into the COPS-PR PIB and, later, delivered to the routers with COPS-PR.

The system database does not have to be modified because it already stores the domain policies such as service level agreements (SLA) with the client traffic demand matrix of the client, the network topology (nodes, link's capacities and costs). The client traffic demand matrix is a maximum traffic configuration which is required in order to assure client contract traffic.

The link weights are transferred to the PIB for their distribution through COPS-PR to the routers or Policy Enforcement Physical Entity (PE-PE) (the same way as the SLS policies). Therefore, the relation of routers of the domain permits the installation of the network policies to the routers.

The COPS-PR Policy Information Base (PIB) has been modified so that the router link weights are sent as a piece of the policies. COPS-PR allows unrequested sending of policies from the PD-FE to the PE-FE and, moreover, it only allows sending parts of the PIB for efficiency improvements. Therefore, whenever it is necessary to change the link weights, only a small part of the PIB is sent through COPS-PR, this being only the modified links weights. Until now, there has been a unique PIB for all core routers, but from now on a PIB for each router is necessary, even though a great part of it will be common to the rest of the core routers (SLS policies).

The PE-PE module installed in every GNU/Linux router fully supports COPS-PR over IPv6 and it is responsible for the configuration of the policies in the PE-PE's PIB and the link weights in a computer running GNU/Linux as a router. The GNU/Linux kernel was already configured for IPv6 support with all the QoS functionalities available to be used in IPv4 as well as IPv6. The load balancing in the GNU/Linux routers is automatically activated if the routing table has multiple paths to a destination using per-destination load balancing.

Whether the use of COPS-PR for the delivery of the links' weights is the most optimal way or not of doing it is a point of discussion. However, the fact the configuration commands also had to be sent in order to configure the new weights when sending the links weights with SSH in the Cisco System routers is an important consideration. With COPS-PR, once the PIB's changes have been implemented, only the small part of the PIB referring to the weight of the concrete changing interfaces are sent.

Thus far, it has been proposed a traffic engineering method to be applied under ITU-T's NGN-GSI specifications and it has been proven its viability by using COPS-PR for the router link weights delivery. It has been successfully optimized weights of a single GNU/Linux router applying the offline algorithm successfully proven in simulations and using the NGN testbed successfully deployed in chapter 4. Despite this headway progress, it has not been optimized and evaluated a complete system with this new proposal. As mentioned in a previous chapter the Networking Research Group testbed has been definitively unmounted and no further testing will be done in it.

6.7. Chapter summary

This chapter has presented the successful implementation of traffic engineering in an IPv6 domain by means of routing optimization. This optimization has been done through the application a HGA algorithm to offline routing with the aim of modifying the OSPF link weights and therefore minimizing the maximum utilization of the domain's links.

Some authors have presented solutions for this problem, but each one has used different parameters and assumptions. An inedited comparative study of the various proposals to solve the OSPFWS problem has been carried out in order to quantify the advantages of the different proposals and it has been demonstrated that HGA algorithms provide good solutions to the complex problem, as opposed to those provided by GA algorithms. Among those, the one which provided marginally better results was the Buriol et al. proposal and which was hence selected to be implemented in the testbed.

The QoS broker has been enhanced with a JAVA application for computing the link weights through the elected HGA algorithm and sending them to the domain's routers, where they are configured automatically. OSPFv3 is the routing protocol which decides the routing inside the autonomous system implemented with IPv6 and computed with the genetically optimized weights. Therefore, this protocol has also been run offline in the QoS broker so the HGA can be applied in order to obtain the new link weights. An initial testbed implementation has been done with commercial Cisco Systems routers. In this first test, weights have been delivered with the SSH protocol.

The ITU-T MS/NGN architecture must provide support for the QoS of the multimedia sessions. This support includes the QoS negotiation and admission, as well as resource control for different end-to-end QoS models. Therefore, it seems reasonable to integrate traffic engineering for intra-domain optimization into this architecture. Working from this hypothesis, this chapter has presented a second testbed implementation of a traffic engineering proposal in ITU-T's NGN-GSI environment. This testbed implementation has been done with GNU/Linux routers running OPSFv3 and with a centralized QoS broker device acting as a decision server (RACF entity). The sending of the optimized link weights has been carried out with the COPS-PR protocol, proposed by the NGN-GSI architecture to manage internal domain policies with PBNM.

These contributions to the thesis have been presented in [207], [208], [222] and [225].

7. CONCLUSIONS AND FUTURE WORK

7.1. Conclusions

This thesis has investigated Quality of Service control architectures as methods for providing seamless end-to-end QoS in the environment of the Next Generation Networks. The basic concept of these architectures are not new but now the ITU-T have provided a global system which is so generic that it can be adapted to any access network, besides the core networks. These QoS control architectures are basic for operators' future growth as they will augment operators' revenues through new enhanced premium quality services while, at the same time, increase the Quality of Experience of users.

From the outset, the IPv6 protocol has been the common base for this thesis, along with the QoS. Therefore, the first objective of the thesis proposed a thorough analysis of the state of the art of the IPv6 conformance and interoperability testing which ended with the empirical study of the IPv6 protocol implementations in end node systems. The IPv6 conformance tests proved that IPv6 was ready to be used in a GNU/Linux testbed with performance guarantees according to specifications. This line of research provided the first contact with the scientific community and the knowledge for writing scientific papers. It also provided the empirical knowledge for conducting testing over real devices and for managing IPv6 supported real GNU/Linux devices, and its associated real life disadvantages, which finally materialized in an IPv6 GNU/Linux testbed from which more research would be conducted.

The implementation of a bandwidth broker architecture responded to the second objective of the thesis. This implementation provided the opportunity to deal with the policy-based network management in order to automate the QoS policy delivery from a centralized device, the BBv6, to the DiffServ nodes. The main drawback while implementing the BBv6 was the incapacity of the interdomain communication protocol, SIBBS, to work with IPv6 addresses which was solved by developing SIBBv6, an enhanced version for IPv6 support that also solved some other flaws of the original protocol. Furthermore, in order to validate the proper functioning of the system, the IPv6 tested was enhanced to fully support the DiffServ specifications and the corresponding PEP implementation of the COPS-PR protocol were then applied to the nodes.

The ITU-T had been developing an end-to-end QoS control generic architecture which fitted the BBv6 architecture because of the use of COPS-PR as the intra-domain communication protocol. By that time the early specifications for the inter-domain interface suggested the use of a dynamic service negotiation protocol so it was decided to use the COPS-SLS protocol after analysing different options. The decision to migrate to a QoS broker architecture based on the ITU-T's NGN-GSI recommendations was then taken. The validation of the architecture was conducted with exactly the same criteria as the first test for the BBv6 architecture and, therefore, provided similar results as the only modification which had been made were in the inter-domain peering between QoS brokers.

Current advances include the decision of the ITU-T to finally recommend a variation of the DIAMETER protocol QoS oriented which is still to be developed. Nevertheless, the DIAMETER protocol was natively designed for authentication, authorization and accounting (AAA) which implies that will be a good option for non-trusted inter-domain communications between different operators but it will probably be too heavy for trusted peering inside the same operator. Therefore a second hybrid protocol architecture proposal has been provided in this thesis, including DIAMETER for non-trusted peering and COPS-SLS for trusted peering.

The knowledge related to the ITU-T end-to-end QoS control architecture, which was acquired throughout this entire process, was applied to develop a proposal for the integration of the QoS management in the new broadband PLC access networks proposals within the ITU-T architecture. The first proposal of QoS control architecture included the COPS-PR's PIB and was made within the framework of the OPERA project specification which opted to be the new IEEE P1901 standard for PLC. Unfortunately, another competitor specification has finally been chosen which has been named the IEEE P1901 draft standard specification. Therefore a second proposal for QoS control architecture has also been provided for this specification. Furthermore the QoS communications requirements of the future Smart Grids have been analysed and the great QoS mappings needs of this technology have been demonstrated. Not only will these mappings be a must inside the own Smart Grids between PLC networks and wireless technologies but they will be also basic for interaction with exterior networks.

The last objective of this thesis was to provide a solution for optimizing routing in NGNs with a genetic algorithm. This algorithm has been implemented in the QoS broker and, therefore, it is currently able to apply traffic engineering to the domain under management by genetically modifying the link weights in order to improve the overall network efficiency. A local search has been added to the genetic algorithm to conform a HGA which has improved the initial performance in quality and speed. With this new heuristic procedure the QoS broker is ready to face real-world problems, self adapting itself according to the information gathered during the evolution and new traffic patterns. The validation of the proposal has been made over a testbed with commercial Cisco Systems routers and also over the GNU/Linux testbed.

The other common nexus along the thesis has been the practical view of the research. This thesis has focused on implementing rather than simulating, even though simulations have been provided when necessary. Nevertheless, the great difficulty of the hand-on testbed versus simulation implementations cannot be underestimated. Moreover the fact that we built our testbed over physical machines provided an additional difficulty over the emulated solution. The results show the viability of our proposals in a real environment and using open systems.

Most of the results presented have focused on providing functionality rather than strict performance results. In order to obtain these detailed performance results with direct application on the operators' networks some of the tests should have been done over high level commercial equipment used in operator's core and aggregation networks, but the economic restrictions and restricted access to the code has pushed towards GNU/Linux solutions on the research path.

As it has become evident over the length this thesis, this is not a regular thesis with a unique classical research line. This thesis has been worked on for six years since 2004 and, even though it has a common story line, it has been providing proposals for different research lines. All the work in this thesis has been realized in the Research Group in Networking Technologies and it is aligned with the research projects this group has taken part in. The diagram in Figure 86 resumes the author's contributions since 2004 in a timeline and relates them to the research lines and to the research projects in which he has participated.

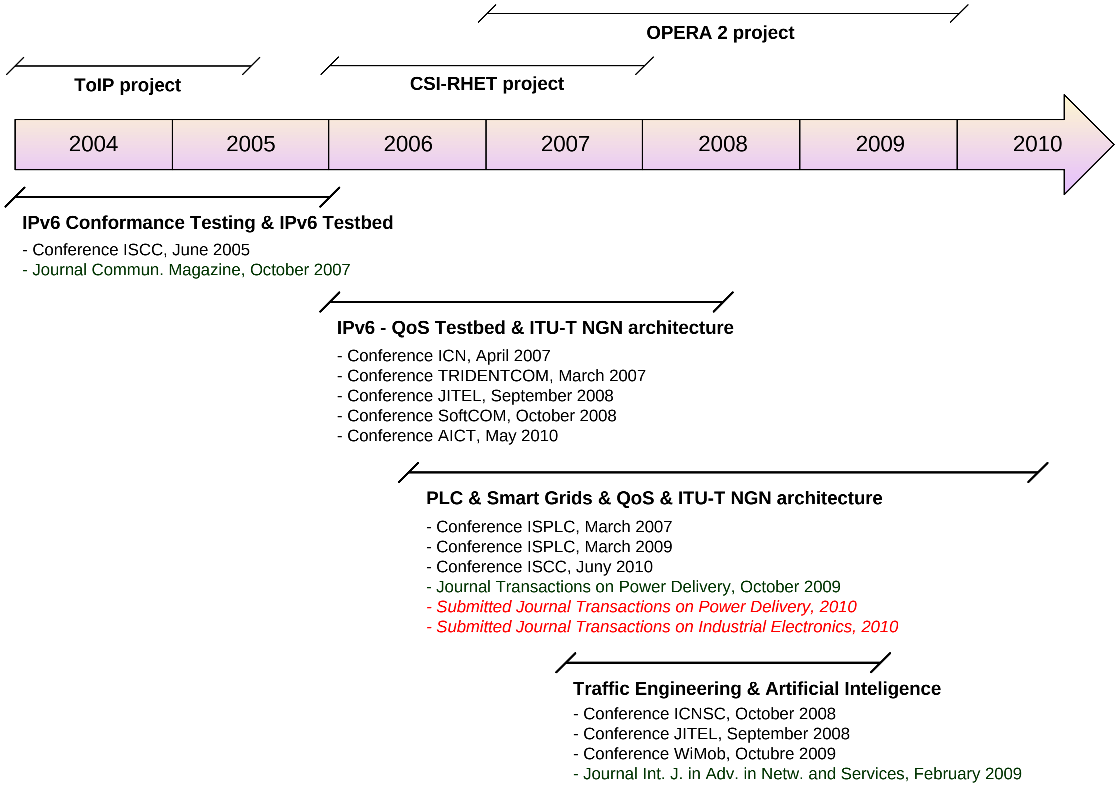

OPERA 2 project
ToIP project
CSI-RHET project
2004
2005
2006
2007
2008
2009
2010
IPv6 Conformance Testing & IPv6 Testbed
- Conference ISCC, June 2005
- Journal Commun. Magazine, October 2007
IPv6 - QoS Testbed & ITU-T NGN architecture
- Conference ICN, April 2007
- Conference TRIDENTCOM, March 2007
- Conference JITEL, September 2008
- Conference SoftCOM, October 2008
- Conference AICT, May 2010
PLC & Smart Grids & QoS & ITU-T NGN architecture
- Conference ISPLC, March 2007
- Conference ISPLC, March 2009
- Conference ISCC, Juny 2010
- Journal Transactions on Power Delivery, October 2009
- Submitted Journal Transactions on Power Delivery, 2010
- Submitted Journal Transactions on Industrial Electronics, 2010
Traffic Engineering & Artificial Inteligence
- Conference ICNSC, October 2008
- Conference JITEL, September 2008
- Conference WiMob, Octubre 2009
- Journal Int. J. in Adv. in Netw. and Services, February 2009

Figure 86: Timeline diagram of the author's contributions and the research lines

All the results, the conclusions, and the lessons learned in this work have also served to fix new objectives that will be approached in further work. In the future work section, a glance ahead is taken and defined future research lines that derive from this thesis are explained.

7.2. Summary of publications

This section lists the author's scientific research productivity. The list of papers in relation to the main goal of this thesis, contributions to the QoS control for the Next Generation Networks, are listed by date:

Journals

- A. Vallejo, A. Zaballos, J.M. Selga, J. Dalmau, "Next Generation QoS control architectures for Broadband Power Lines Networks", Submitted to IEEE Transactions on Power Delivery Journal, 2010.
- A. Zaballos, A. Vallejo, M. Majoral, J.M. Selga, "A new paradigm for smart grids communications architecture", Submitted to IEEE Transactions on Industrial Electronics, 2010.
- A. Zaballos, A. Vallejo, M. Majoral, J.M. Selga, "Survey and Performance Comparison of AMR over PLC standards", Power Delivery, IEEE Transactions on, vol.24, no.2, pp.604-613, April 2009.
- A. Vallejo, A. Zaballos, D. Vernet, A. Orriols-Puig, J. Dalmau, "A Traffic Engineering proposal for ITU-T NGNs using Hybrid Genetic Algorithms", Advances in Networks and Services, International Journal of, vol 2, no 1, February 2009.
- A. Vallejo, J. Ruiz, A. Zaballos, J. Abella, J.M. Selga, “State of the art of IPv6 conformance and interoperability testing”, Communications Magazine, IEEE, vol.45, no.10, pp.140-146, October 2007.

Conferences

- A. Vallejo, A. Zaballos, A. Campos, J. Dalmau, "Optimizing the usage of COPS protocol in ITU-T NGN architecture", AICT, May 2010. In press.
- A. Zaballos, A. Vallejo, J.J. Jimenez, J.M. Selga, "QoS Broker based management for heterogeneous Smart Electricity Networks", IEEE Symposium on Computers and Communications (ISCC'10), June 2010. In press.
- A. Zaballos, A. Vallejo, P. Terradellas, J.M. Selga, "A genetic-based QoS aware routing for ubiquitous sensor networks", Wireless and Mobile Computing, Networking and Communications, 2009. WIMOB 2009. IEEE International Conference on, vol., no., pp.129-134, October 2009.
- A. Zaballos, A. Vallejo, J.J. Jimenez, J.M. Selga, "QoS Broker based architecture design for the PLC access network", Power Line Communications and Its Applications, 2009. ISPLC 2009. IEEE International Symposium on, vol., no., pp.205-210, April 2009.
- A. Vallejo, A. Zaballos, D. Vernet, D. Cutiller, J. Dalmau, "Implementation of Traffic Engineering in NGNs using Hybrid Genetic Algorithms", Systems and Networks Communications, 2008. ICSNC '08. 3rd International Conference on, vol., no., pp.262-267, October 2008.
- A. Vallejo, A. Zaballos, X. Canaleta, J. Dalmau, "Estudio de la gestión de la QoS extremo-extremo en la arquitectura ITU-T IMS/NGN", VII Jornadas de Ingeniería Telemática, JITEL 2008, September 2008.
- A. Zaballos, A. Vallejo, J.M. Selga, X. Canaleta, "Aplicación de AGs en el encaminamiento con QoS en redes USN Access Networks", VII Jornadas de Ingeniería Telemática, JITEL 2008, September 2008.
- A. Vallejo, A. Zaballos, X. Canaleta, J. Dalmau, "End-to-End QoS Management Proposal for the ITU-T IMS/NGN Architecture", Software, Telecommunications and Computer Networks, 2008. SoftCOM 2008. 16th International Conference on, vol., no., pp.147-151, September 2008.
- A. Vallejo, A. Zaballos, J. Abella, G. Villegas, J.M. Selga, “Evaluation of a Policy-Based QoS Management Architecture over an IPv6 DiffServ testbed”, IEEE International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TridentCom '07), 3rd Tridentcom 2007, May 2007.
- A. Vallejo, A. Zaballos, J. Abella, J.M. Selga, C. Duz, “Performance of a policy-based management system in IPv6 networks using COPS-PR”, International Conference on Networking (ICN'07), 6th ICN 2007, April 2007.

- A. Zaballos, A. Vallejo, G. Ravera, J.M. Selga, "Simulation and modeling of the coexistence of polling and contention in PLC based AMR systems", Power Line Communications and Its Applications, 2007. ISPLC 2007. IEEE International Symposium on, vol., no., pp.110-115, March 2007.
- J. Ruiz, A. Vallejo, J. Abella, “IPv6 Conformance and Interoperability Testing”, IEEE Symposium on Computers and Communications (ISCC'05), 10th ISCC, pp. 83-88, June 2005.

Other author’s scientific research productivity not directly related to this thesis is also listed:

- A. Zaballos, A. Vallejo, P. Terradellas, J.M. Selga, "Issues of QoS multipath routing protocol for SEN’s data networks", AICT, May 2010. In press.
- X. Canaleta, P. Ros, A. Vallejo, D. Vernet, A. Zaballos, "A system to extract social networks based on the processing of information obtained from Internet", Proceedings of CCIA, October 2008.
- A. Zaballos, A. Vallejo, J. Vives, J. Abella, "Modeling and Analyzing the Star of Chains Topology for Wireless Sensor Networks", Proceedings of Opnetwork’2008, August 2008.
- A. Zaballos, A. Vallejo, G. Corral, J. Abella, “AdHoc routing performance study using OPNET Modeler", Proceedings of Opnetwork’2006, August 2006.

7.3. Future work

In the previous section, the contributions made in the various lines of research have been described. The QoS is a key factor for the success of the next generation Internet in the forthcoming years. Increasing traffic demands and the expected QoS will challenge the whole IP communication system. MPLS Traffic Engineering (MPLS-TE) is a growing implementation in today's service provider networks. MPLS adoption in service provider networks has increased manifold due to its inherent TE capabilities. MPLS TE allows the MPLS-enabled network to replicate and expand upon the TE capabilities of the outdated layer-2 ATM and Frame Relay networks. MPLS uses the reachability information provided by layer-3 routing protocols and operates like a layer-2 ATM network. Traffic Engineering via MPLS improves the reliability of service provider networks in two ways. First, it allows service providers to route traffic around congested points and to avoid "hot spots” in the network. Secondly, the MPLS paths that traverse a network can be set up to be both redundant and load sharing. This will allow service providers to assure that critical traffic always has a path through the network.

MPLS uses RSVP-TE (traffic engineering) and Constraint-based Label-Distribution Protocol (CR-LDP) for special-purpose signalling, even though the IETF has stated in RFC-3468 that “MPLS Working Group consensus to undertake no new work on CR-LDP and focus on RSVP-TE as signalling protocol for traffic engineering applications for MPLS” [5].

There is also a lot of work going on in the DiffServ aware MPLS-TE (MPLS DS-TE) integration area. This appears to be the only viable approach to the scalability problem that ISPs, operators and carriers face when dealing with flows and service classes in core networks. MPLS-DS-TE networks are beginning to be deployed and are far more complex than best-effort networks, which is why it is not possible to think in manual operations to achieve Traffic Engineering. Besides, its optimization makes the use of centralized servers like the already mentioned bandwidth brokers and PCE a necessary requirement. The PCE architecture for MPLS-TE and MPLS DS-TE networks has been proposed in order to manage LSP path from a centralized server. These servers must solve multiple aspects of those networks among which there are at least the creation and elimination LSPs, the routing of LSPs in real time, the global analysis and optimization of the layout of LSPs, their dimensioning, the adaptation to traffic changes, the prioritization of classes of service and others. Because this is an important and current research area, the migration of the QoS broker device to a PCE device is proposed in compliance with the directives given by the IETF PCE working group [97] for the automation and optimization of MPLS-DS-TE core networks with network intelligence and QoS guarantees, evaluating the different possible of architectures. Even though the functions of the QoS broker and the PCE are different, but both devices manage the introdomain nodes, use COPS-PR for policy delivery and need to establish inter-domain peering with neighbouring devices.

These PCE different solutions could be developed for the computation and optimization of MPLS networks, even though it proposed to focus the work basically in the application of techniques based on Artificial Intelligence. The LSPs must be set in real time which creates a global vision problem. Therefore two levels of routing are needed, one for the establishment of real-time LSP, which might be distributed or centralized according to timing requirements, and another one for making global calculations with the overall PCE available information. This second level should be offline routing and provides the network global optimization vision. The offline routing might be applied periodically or whenever necessary and could incorporate genetic algorithms. The HGA algorithm presented in chapter 6 can be easily adapted for this new function and therefore preliminary simulations on this field can be found in Appendix F [209].

Several implementations of MPLS for GNU/Linux exist, even though not many support MPLS-TE and, even less, MPLS DS-TE. In fact there are no complete GNU/Linux implementations of MPLS DS-TE which are non-commercial, basically due to problems with the RSVP-TE signalling daemons. See Appendix E for more information on the state of the art of MPLS-DS-TE implementations.

Moreover, the recent emergence of GNS3, a free powerful network simulator based on Cisco IOS, which allows running Cisco IOS on regular computers, will also feature prominently in the and therefore the use of computers as Cisco routers. This project is an open source, free program that may be used on multiple operating systems, including Windows, Linux, and MacOS X. This software avoids great expense on high level Cisco Systems hardware in order to test core routers performance and opens a new line for practical research.

BIBLIOGRAPHY

[1] Agilent Technologies, Agilent N2X, May 2010, [Online]. Available: www.agilent.com

[2] H. Alvestrand, "A Mission Statement for the IETF", IETF RFC-3935, October 2004.

[3] Ambient Networks project, "Connecting Ambient Networks - Architecture and Protocol Design (Release 1)." Del. D 3.2, mar. 2005.

[4] AQUILA project, May 2010, [Online]. Available: www-st.inf.tu-dresden.de/aquila/

[5] L. Andersson, G. Swallow, "The Multiprotocol Label Switching (MPLS) Working Group decision on MPLS signaling protocols ", IETF RFC-3468, February 2003.

[6] D.V. Andrade,_L.S. Buriol, M. Resende, M. Thorup,_Survivable composite-link IP network design with OSPF routing, The Eighth INFORMS Telecommunications Conference, Dallas, Texas, April 2006.

[7] Atlantis, "RSVP-TE daemon for diffserv over MPLS under Linux", May 2010, [Online]. Available: http://dsmpls.atlantis.ugent.be/

[8] M. Bamatraf, M. Othman, "Improved balancing heuristics for optimizing shortest path routing", Computer Communications. Vol. 30, pp. 1513-1526, May 2007.

[9] A. Banchs, A. Azcorra, C. Garcia, R. Cuevas, "Applications and challenges of the 802.11e EDCA mechanism: an experimental study," Network, IEEE, vol.19, no.4, pp. 52-58, July-Aug. 2005.

[10] L. Berry, S. Koehler, D. Staehle, P. Tran-Gia, "Fast heuristics for optimal routing in large IP networks", Technical Report 262, University of Wuerzburg, July 2000.

[11] BII Group, BII laboratory, May 2010, [Online]. Available: www.biigroup.com

[12] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An Architecture for Differentiated Service", IETF RFC-2475, December 1998.

[13] R. Bless, “Dynamic Aggregation of Reservations for Internet Services”, Telecommunication Systems, Vol. 6, No. 1, pp. 33-52, November 2004.

[14] R. Braden, Ed., L. Zhang, S. Berson, S. Herzog, S. Jamin, “Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification ”, IETF RFC-2205, September 1997.

[15] S. Bradner, A. Mankin, "The Recommendation for the IP Next Generation Protocol", IETF RFC-1752, January 1995.

[16] S. Bradner, “The Internet Standards Process - Revision 3”, IETF RFC-2026, October 1996.

[17] L.S. Buriol, M. Resende, C. Ribeiro, M. Thorup. “A Hybrid Genetic Algorithm for the Weight Setting Problem in OSPF/IS-IS Routing”, Networks, Vol. 46, pp. 36-56. November 2005.

[18] CADENUS project, May 2010, [Online]. Available: www.ist-world.org/ProjectDetails.aspx?ProjectId=08e5f7c63d044d4a956f0ca71ec8d03a

[19] P. Calhoun, J. Loughney, E. Guttman, G. Zorn, J. Arkko, "Diameter Base Protocol", IETF RFC3588, September 2003.

[20] R. Callon, “Use of OSI IS-IS for Routing in TCP/IP and Dual Environments”, IETF RFC-1195, December 1990.

[21] J. Case, M. Fedor, M. Schoffstall and J. Davin, "Simple Network Management Protocol (SNMP)", IETF RFC 1157, may. 1990.

[22] V. Cerf, Y. Dalal, C. Sunshine, ”Specification of Internet Transmission Control Program”, IETF RFC-675, December 1974.

[23] K. Chan, et al., “COPS Usage for Policy Provisioning (COPS-PR)”, IETF RFC-3084, March 2001.

[24] K. Chan, R. Sahita, S. Hahn, K. McCloghrie, “Differentiated Services Quality of Service Policy Information Base”, IETF RFC-3317, March 2003.

[25] J. Chen, A. McAuley, V. Sarangan, S. Baba, and Y. Ohba, “Dynamic Service Negotiation Protocol (DSNP) and wireless DiffServ”, IEEE International Conference on Communications (ICC’02), 37th ICC, Vol. 2, April 2002.

[26] China education and research network, The Chinese Next-generation Internet Demonstration (CNGI) Project, May 2010, [Online]. Available: www.edu.cn

[27] Cisco Systems, “Configuring OSPF”, Cisco Systems, Inc., San Jose, USA, August, 2006.

[28] Cisco Systems, “Implementing OSPF for IPv6”, Cisco Systems, Inc., San Jose, USA, August, 2007.

[29] R. Coltun, D. Ferguson, J. Moy, A. Lindem, “OSPF for IPv6”, IETF RFC-5340, July 2008.

[30] R. Coltun, T. Li, H. Smit, “OSPF for IPv6”, IETF RFC-RFC5305, October 2008.

[31] A. Conta, S. Deering, "Internet Control Message Protocol (ICMPv6) for the IPv6 Specification", IETF RFC-2463, December 1998.

[32] A. Conta, S. Deering, M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the IPv6 Specification", IETF RFC-4443, March 2006.

[33] Corinex Communications Corp., "Broadband over Powerline for Smart Grids Technology Brief", May 2010 [Online]. Available: www.corinex.com/images/tb/

[34] L.A. DaSilva, "QoS mapping along the protocol stack: discussion and preliminary results," Proc. 2000 IEEE Int. Conf. on Commun. (ICC 2000), vol.2, pp. 713–17, June 2000.

[35] B. Davie, et.al., “An Expedited Forwarding PHB (Per-Hop Behavior)”, IETF RFC-3246, March 2002.

[36] S. Deering, R. Hinden, ”Internet Protocol, Version 6 (IPv6) Specification”, IETF RFC-1883, December, 1995.

[37] S. Deering, R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification”, IETF RFC-2460, December 1998.

[38] Delivery of Advanced Network Technology to Europe (DANTE) Limited, GÉANT2 Project, May 2010, [Online]. Available: www.geant2.net

[39] E. Dijkstra, “A note on two problems in connection of graphs”, Numerical mathematics, Vol. 1, pp. 269-271, 1959.

[40] DS2, “DS2: technology: How it works,” May 2010, [Online]. Available: www.ds2.es

[41] Z. Duan, Z. Zhang, Y. Thomas Hou, L. Gao, “A core stateless bandwidth broker architecture for scalable support of guaranteed services, Parallel and Distributed Systems”, IEEE Transactions on parallel and distributed Systems, Vol. 15, No. 2, pp. 167-182, February 2004.

[42] D. Durham, Ed., J. Boyle, R. Cohen, S. Herzog, R. Rajan, A. Sastry, “The COPS (Common Open Policy Service) Protocol”, IETF RFC-2748, January 2000.

[43] D27: First draft of the Opera Specification Version 2, OPERA project, May 2010, [Online]. Available: www.ist-opera.org

[44] EPRI, Electric Power Research Institute, IntelliGrid initiative, May 2010, [Online] Available: http://intelligrid.epri.com

[45] ENIGMA project, May 2010, [Online]. Available: www.i2cat.net/es/projecte/e3ms

[46] ENTHRONE project, May 2010, [Online]. Available: www.ist-enthrone.org

[47] M. Ericsson, M.G.C. Resende, P.M. Pardalos, “A Genetic Algorithm For The Weight Setting Problem in OSPF Routing”, Journal of Combinatorial Optimization, Vol. 6, N. 3, pp. 299-333, September 2002.

[48] EUQoS project, May 2010, [Online]. Available: www.euqos.eu

[49] European Commission, eEurope Action Plan, May 2010, [Online]. Available: www.e-europestandards.org

[50] European Comission, “Strategic Research Agenda for Europe’s electricity networks of the future,” May 2010. [Online]. Available: www.smartgrids.eu

[51] European IPv6TF, May 2010, [Online]. Available: www.IPv6tf.org

[52] European SmartGrids Technology Platform, "Strategic Deployment Document for Europe’s Electricity Networks of the Future -SmartGrids_SDD_Draft_25_sept_2008", May 2010 [Online]. Available: http://www.smartgrids.eu/documents/3rdGA/

[53] ETSI, European Telecommunications Standards Institute, May 2010, [Online]. Available: www.etsi.org

[54] ETSI, European Telecommunications Standards Institute, ES 201 873-1 to 9 v.3.1.1., “The Testing and Test Control Notation version 3, Part 1 to 9”, Sophia-Antipolis, 2005.

[55] ETSI, European Telecommunications Standards Institute, ES 282 003 V1.1.1 (2006-03), “Resource and Admission Control Sub-system (RACS); Functional Architecture.” Sophia-Antipolis, 2007.

[56] ETSI, European Telecommunications Standards Institute, ETS 300-406, "Methods for Testing and Specification (MTS), Protocol and profile conformance testing specifications; Standardization methodology", Sophia-Antipolis, 1995.

[57] ETSI, European Telecommunications Standards Institute, IPv6 plugtests, May 2010, [Online]. Available: www.etsi.org/plugtests

[58] ETSI, European Telecommunications Standards Institute, IPv6 Testing Project, May 2010, [Online]. Available: www.ipt.etsi.org

[59] ETSI, European Telecommunications Standards Institute, ETSI NGN, May 2010 [Online]. Available: www.etsi.org/Technologies/NextGenerationNetworks.aspx

[60] ETSI, European Telecommunications Standards Institute, TS 102 351 v.1.1.1., “TTCN-3 IPv6 Test Specification Toolkit”, Sophia-Antipolis, 2004.

[61] ETSI, European Telecommunications Standards Institute, TS 102 514 to 517, “IPv6 Core: Requirements Catalogue, Conformance TSS & TP, Conformance Test Suite, Interoperability Test Suite”, Sophia-Antipolis, 2006.

[62] ETSI, European Telecommunications Standards Institute, TS 102-237-1, "Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 4, Interoperability test methods and approaches, Part 1: Generic approach to interoperability testing”, Sophia-Antipolis, 2005.

[63] F. Le Faucheur, “Protocol Extensions for Support of DiffServ-aware MPLS Traffic Engineering”, IETF RFC-4124, June 2005.

[64] M. Feng, R. Leung, A. Do-Sung Jun, “Summer Report 1999 Linux Network”, Universisty of Toronto, 1999.

[65] FERC, Federal Energy Regulation Commission, “Smart Grid Policy,” 2009. [Online]. Available: www.ferc.gov/whatsnew/comm-meet/2009/071609/E-3.pdf

[66] B. Fink, 6bone backbone, May 2010, [Online]. Available: www.6bone.net

[67] B. Fortz, M. Thorup, “Internet traffic engineering by optimizing OSPF weights”, IEEE IEEE International Conference on Computer Communication (INFOCOM’00), 19th INFOCOM, Vol. 2, pp. 519-528, March 2000.

[68] B. Fortz, J. Rexford, M. Thorup, “Traffic Engineering with Traditional IP Routing Protocols”, IEEE Communications Magazine, Vol.40, No.10, pp.118-124, October 2002.

[69] B. Fortz, M. Thorup, “Optimizing OSPF/IS-IS Weights in a Changing World”, IEEE Journal of Selected Areas Communication v20.756-767, November 2002.

[70] B. Fortz, M. Thorup, “Increasing Internet capacity using local search”, Computacional optimization and applications, Vol. 29, N. 1, pp. 13-48, October 2004.

[71] Gigle Semiconductor, “Home multimedia networks demand higher performance. A white paper”, May 2010, [Online]. Available: www.gigle.biz

[72] D.E. Goldberg, “Genetic Algorithms in Search, Optimization & Machine Learning”, Addison-Wesley, Massachusetts, 1989.

[73] D.E. Goldberg, “The Design of Innovation: Lessons from and for Competent Genetic Algorithms”, Kluwer Academic Publishers, 2002.

[74] Google Inc., Google powermeter, May 2010, [On-line]. Available: www.google.com/powermeter/

[75] D. Grossman, “New Terminology and Clarifications for DiffServ”, IETF RFC-3260, April 2002.

[76] M. Günter, T Braun, I Khalil, “An Architecture for Managing QoS-enabled VPNs over the Internet”, IEEE Conference on Local Computer Networks (LCN’99), 24th LCN, pp.122-131, October 1999.

[77] J. Harrison, J. Berger, M. Bartlett, “IPv6 Traffic Engineering in IS-IS”, IETF Internet Draft “draft-ietf-isis-ipv6-te-04”, August 2007.

[78] M. Hashmani, M. Yoshida, "ENICOM's bandwidth broker", Symposium on Applications and the Internet-Workshops (SAINT’01 Workshops), 3rd SAINT, pp. 213-220, January 2001.

[79] Hasso, “Essay about developing in XORP”, May 2010, [Online]. Available: http://hasso.linux.ee/doku.php/english:network:xorpsucks

[80] J. Heinanen, F. Baker, W. Weiss, J. Wroclawski, “Assured Forwarding PHB Group”, IETF RFC-2597, June 1999.

[81] S. Herzog, Ed, J. Boyle, R. Cohen, D. Durham, R. Rajan, A. Sastry, “COPS usage for RSVP”, IETF RFC-2749, January 2000.

[82] R. Hinden, S. Deering, “IP Version 6 Addressing Architecture”, IETF RFC-4291, February 2006.

[83] J.H. Holland, “Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artiçial Intelligence”. MIT Press/ Bradford Books edition, 1992.

[84] HomePlug Powerline Alliance, “HomePlug AV White Paper”, May 2010, [Online]. Available: www.homeplug.org

[85] C. Hopps, “Routing IPv6 with IS-IS”, IETF RFC-5308, October 2008.

[86] IANA, Internet Assigned Numbers Authority, May 2010, [Online]. Available: www.iana.org

[87] IBM Corp., International Business Machines Corporation, IPv6 enablement at IBM, May 2010, [Online]. Available: www.306.ibm.com/software/os/systemz/ipv6

[88] ICANN, Internet Corporation for Assigned Names and Numbers, May 2010, [Online]. Available: www.iana.org

[89] IEC, International Electrotechnical Commission, Communication networks and systems in substations, IEC Standard 61850, 2003.

[90] IEEE P1901 Working Group, May 2010 [Online]. Available: http://grouper.ieee.org/groups/1901/

[91] IEEE P2030 Working Group, May 2010 [Online]. Available: http://grouper.ieee.org/groups/scc21/2030/2030_index.html

[92] IEEE 802 Committee of the IEEE Computer Society, IEEE P802.11e Amendment to IEEE Std 802.11, “Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: MAC Quality of Service (QoS) Enhancements”, 2005.

[93] IETF, Internet Engineering Task Force, May 2010, [Online]. Available: www.ietf.org

[94] IETF, Internet Engineering Task Force, Diameter Maintenance and Extensions (DIME) Working Group Charter, May 2010, [Online]. Available: www.ietf.org/wg/dime/charter/

[95] IETF, Internet Engineering Task Force, Differentiated Services (DiffServ) Working Group Charter, May 2010, [Online]. Available: www.ietf.org/html.charters/OLD/ diffserv-charter.html

[96] IETF, Internet Engineering Task Force, IPv6 Working Group (IPv6), May 2010, [Online]. Available: www.ietf.org/html.charters/IPv6-charter.html

[97] IETF, Internet Engineering Task Force, Path Computation Element (PCE) Working Group Charter, May 2010, [Online]. Available: www.ietf.org/html.charters/pce-charter.html

[98] IETF, Internet Engineering Task Force, Policy Framework (Policy) Working Group Charter, May 2010, [Online]. Available: www.ietf.org/html.charters/OLD/policy-charter.html

[99] IETF, Internet Engineering Task Force, Resource Allocation Protocol (RAP) Working Group Charter, May 2010, [Online]. Available: www.ietf.org/html.charters/OLD/rap-charter.html

[100] IETF, Internet Engineering Task Force, Configuration Management with SNMP (snmpconf) Working Group Charter, May 2010, [Online]. Available: www.ietf.org/html.charters/OLD/snmpconf-charter.html

[101] Internet Usage Stats and Population Report, China, Internet, May 2010, [Online]. Available: http://www.internetworldstats.com/asia/cn.htm

[102] Internet 2, “QBone Signaling Design Team”, May 2010, [Online]. Available: http://qos.internet2.edu

[103] IPv6 Forum, May 2010, [Online]. Available: www.IPv6forum.org

[104] IPv6 Forum, IPv6 Ready Logo Program, May 2010, [Online]. Available: www.IPv6ready.org

[105] IPv6 Promotion Council, Certification Working Group, May 2010, [Online]. Available: www.v6pc.jp/en/wg/certificationWG

[106] IPv6 Promotion Council, SIP IPv6 Sub-Working Group, IMS IPv6 conformance test, May 2010, [Online]. Available: http://cert.v6pc.jp/ims-ipv6/

[107] IRISA, Institut de Recherche en Informatique et Systèmes Aléatoires, Tests de Conformité & d'Interopérabilité des Protocoles Internet (TIPI), May 2010, [Online]. Available: www.irisa.fr/tipi

[108] K. Ishiguro, V. Manral, A. Davey, "Traffic Engineering Extensions to OSPF version 3", IETF RFC-5329, September 2008.

[109] ISO/IEC 9646-1 to 7, "Information technology - Open systems interconnection - Conformance testing methodology and framework. Part 1 to Part 7", Genève, 1995.

[110] ISOC, Internet Society, May 2010, [Online]. Available: www.isoc.org

[111] ITU-T, Joint Co-ordination Activity on Home Networking (JCA-HN), May 2010 [Online]. Available: www.itu.int/ITU-T/jca/hn/index.phtml

[112] ITU-T, Draft Recommendations Q.3323.x: "Protocol at the interface between Policy Decision Physical Entity (PD-PE) and Policy Enforcement Physical Entity PE-PE (Rw interface) version 2", 2009.

[113] ITU-T, Draft Recommendations Q.3324.x: "Protocol at the interface between Transport Resource Control Physical Entity (TRC-PE) and Transport Phisical Entity (T-PE) (Rc interface) version 2", 2009.

[114] ITU-T, Recommendation Y.1291, "An architectural framework for support of quality of service in packet networks", Genève, 2004.

[115] ITU-T, Recommendation Y.2001, "Overview NGN", Genève, 2004.

[116] ITU-T, Recommendation Y.2012, "Functional Requirements and Architucture of the NGN", Genève, 2006.

[117] ITU-T, Recommendation Y.2021, "IP Multimedia Subsystem for NGN", Genève, 2006.

[118] ITU-T, Recommendation Y.2051, "General overview of IPv6 in NGN", Genève, 2008.

[119] ITU-T, Recommendation Y.2111, "Resource and Admission Control Functions in NGN (version 2)", Genève, 2008.

[120] ITU-T, Recommendation Y.2112, "A QoS control architecture for Ethernet-based IP access networks", Genève, 2007.

[121] ITU-T, Recommendation Z.140 to Z.146, "The Testing and Test Control Notation version 3", Genève, 2006.

[122] ITU-T, Study Group 11, May 2010, [Online]. Available: www.itu.int/ITU-T/studygroups/com11/

[123] ITU-T, Study Group 15, May 2010, [Online]. Available: www.itu.int/ITU-T/studygroups/com15/

[124] IXIA, IxANVL, May 2010, [Online]. Available: www.ixiacom.com

[125] Japanese IPv6 Promotion Council, May 2010, [Online]. Available: www.v6pc.jp

[126] S. Jha, M. Hassan, "Implementing bandwidth broker using COPS-PR in Java", IEEE Conference on Local Computer Networks (LCN'01), 26th LCN, pp. 178-179, November 2001.

[127] Keio University, KAME Project, May 2010, [Online]. Available: www.kame.net

[128] J. Korhonen, Ed., H. Tschofenig, E. Davies, "Quality of Service Parameters for Usage with Diameter", IETF RFC5624, August 2009.

[129] J. Korhonen, H. Tschofenig, M. Arumaithurai, M. Jones, Ed., A. Lior, "Traffic Classification and Quality of Service (QoS) Attributes for Diameter", IETF RFC5777, February 2010.

[130] C. N. Krishnan, "Powerline as Access Medium--A Survey", Dept. of Electronics Engg. M.I.T Campus of. Anna University, Chennai, 2001.

[131] R.J. Linn, M.Ü. Uyar, “Conformance testing methodologies and architectures for OSI protocols”, IEEE Computer Society Press, Los Alamitos, USA, 1995, ISBN: 0-8186-5352-3.

[132] J. Manner, G. Karagiannis, G. Karagiannis, and A. McDonald, “NSLP for Quality-of-Service Signalling,” IETF Internet draft, draft-ietf-nsis-qos-nslp-16.txt, feb. 2008.

[133] M. Mani, N. Crespi: "Inter-Domain QoS Control Mechanism in IMS based Horizontally Converged Networks", International Conference on Networking and Services (ICNS 2007), July 2007, p. 82.

[134] S.I. Maniatis, E.G. Nikolouzou, I.S. Venieris, “QoS Issues in the converged 3G Wireless and Wired Networks”, IEEE Communications Magazine, Vol. 40, No. 8, pp. 44-53, August 2002.

[135] M. Marchese, “QoS Over Heterogeneous Networks”, Wiley, 2007.

[136] V. Marques, et al., "An IP-based QoS architecture for 4G operator scenarios," Wireless Communications, IEEE, vol.10, no.3, pp. 54-62, June 2003.

[137] J. McCann, S. Deering, J. Mogul, "Path MTU Discovery for IP version 6", IETF RFC-1981, August 1996.

[138] K. McCloghrie, et al., “Structure of Policy Provisioning Information (SPPI)”, IETF RFC-3159, August 2001.

[139] MESCAL project, May 2010, [Online]. Available: www.mescal.org/

[140] Microsoft Corporation, Microsoft Technet IPv6, May 2010, [Online]. Available: www.microsoft.com/technet/network/IPv6/default.mspx

[141] S. Moseley, S. Randall, A. Wiles, “Experience within ETSI of the combined roles of conformance testing and interoperability testing”, IEEE Conference on Standardization and Innovation in Information Technology (SIIT’03), 3rd SIIT, pp. 177-189, October 2003.

[142] J. Moy, “OSPF Version 2”, IETF RFC-2328, April 1998.

[143] P. Moscato, “On Evolution, Search, Optimization, Genetic Algorithms and Material Arts: Towards Memetic Algorithms, Caltech Concurrent Computation Program, C3P Report 826, 1989.

[144] P. Moscato, “Optimization of memetic Algorithms”, Caltech Concurrent Computation Program, C3P Report 1024, 1995.

[145] MPLS-Linux, May 2010, [Online]. Available: http://mpls-linux.sourceforge.net/

[146] E. Mulyana and U. Killat, “An Alternative Genetic Algorithm to Optimize OSPF”, In Internet Traffic Engineering and Traffic Management, 15-th ITC Specialist Seminar, Wuerzburg Germany, July 2002.

[147] E. Mulyana and U. Killat, “A Hybrid Genetic Algorithm Approach for OSPF Weight Setting Problem”, 2nd Polish-German Teletraffic Symposium PGTS, September 2003.

[148] P. Nanda, A.J. Simmonds, K. Rajput: “Policy Based Network Architecture in Support for Guaranteed QoS”, International Conference on Information Technology: Prospects and Challenges (ITPC’03), 2nd ITPC, May 2003.

[149] T. Narten, E. Nordmark, W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)”, IETF RFC-4861, September 2007.

[150] NAv6TF, May 2010, [Online]. Available: www.nav6tf.org

[151] R. Neilson, J. Wheeler, F. Reichmeyer, S. Hares, “A Discussion of bandwidth broker Requirements for Internet2 Qbone Deployment”, v.0.7, Internet2 Qbone BB Advisory Council, August 1999.

[152] Net-O2 Technologies, Net-O2 Attest. May 2010, [Online]. Available: www.net-o2.com

[153] T. M. T. Nguyen, N. Boukhatem, Y.G. Doudane, G. Pujolle, “COPS-SLS: A Service Level Negotiation Protocol for the Internet”, IEEE Communications Magazine, Vol. 40, No. 5, pp. 158-165, May 2002.

[154] T.M.T. Nguyen, N. Boukhatem, G. Puiolle, "COPS-SLS usage for dynamic policy-based QoS management over heterogeneous IP networks", IEEE Network, Vol. 17, N. 3, pp. 44-50, may. 2003.

[155] K. Nichols, S. Blake, F. Baker, D. Black, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers”, IETF RFC-2474, December 1998.

[156] K. Nichols, V. Jacobson, L. Zhang, “A Two-Bit Differentiated Services Architecture for the Internet”, IETF RFC-2638, July 1999.

[157] NIST, National Institute of Standards and Technology, “A Profile for IPv6 in the U.S. Government ver. 1.0”, July 2008.

[158] NIST, National Institute of Standards and Technology, "NIST Framework and Roadmap for Smart Grid Interoperability Standards Release 1.0 (Draft)", September 2009.

[159] NIST, National Institute of Standards and Technology, “Smart Grid Priority Action Plan Guidelines for the Use of IP Protocol Suite for Smart Grid”, NIST draft publication, 2009.

[160] OPERA project, May 2010, [Online]. Available: www.ist-opera.org

[161] OSPF-API, May 2010, [Online]. Available: http://wiki.quagga.net

[162] P. Pan, E. Hahne, and H. Schulzrinne, “BGRP: A Tree-Based Aggregation Protocol for interdomain Signaling”, Journal of Communications and Networks, Vol. 2, No. 2, pp. 157-167, June 2000.

[163] C.K. Park, et al., "Definition of Common PLC MIB and Design of MIB Mapper for Multi-vendor PLC Network Management", Power Line Communications and Its Applications, 2009. ISPLC 2009. IEEE International Symposium on, vol., no., 2009.

[164] T. Pereira, A. Beller, E. Jamhour, M. Fonseca, “QoS Management on MobileIP networks using COPS-PR”, Network Operations and Management Symposium 2006 (NOMS’06), 10th IEEE/IFIP NOMS, pp. 1-4, April 2006.

[165] J. Postel, “Internet Protocol”, IETF RFC-791, September 1981.

[166] PRIME project, May 2010, [Online]. Available: www.prime-alliance.org

[167] J. K. Pritchard, “Energy Security: Reducing Vulnerabilities to Global Energy Networks,” May 2009, [Online]. Available: www.dtic.mil

[168] Quagga Routing Suite, May 2010, [Online]. Available: www.quagga.net

[169] F. Rafique, “Implementing Restoration Routing in a LINUX Router”, Lahore University Ed., May 2005.

[170] D. Rawlins, A. Kulkarni, M. Bokaemper, K. Chan, “Framework for Policy Usage Feedback for Common Open Policy Service with Policy Provisioning (COPS-PR)”, IETF RFC-3483, March 2003.

[171] W. Reichl, E.-O. Ruhle, “NGA, IP-Interconnection and their Impact on Business Models and Competition”, in: Communications & Strategies, no. 69, pp. 41-62, February 2008.

[172] M. Resende, et al., "A memetic algorithm for the weight setting problem in OSPF routing", INFORMS Telecommunications Conference (INFORMS’03), 6th INFORMS, March 2003.

[173] A. Riedl, “A Hybrid Genetic Algorithm for Routing Optimization in IP Networks Utilizing Bandwidth and Delay Metrics”, IEEE Workshop on IP Operations and Management (IPOM’02), 2nd IPOM, October 2002.

[174] A. Riedl, D.A. Schupke, “A Flow-Based Approach for IP Traffic Engineering Utilizing Routing Protocols With Multiple Metric Types”, INFORMS Telecommunications Conference (INFORMS’03), 6th INFORMS, March 2003.

[175] A. Riedl, D. A. Schupke, “Routing optimization in IP networks utilizing additive and concave metrics”, IEEE/ACM Transactions on Networking, Vol. 15, N. 5, pp. 1136-1148, 2007.

[176] C.E. Rothenberg and A. Roos, “A Review of Policy-Based Resource and Admission Control Functions in Evolving Access and Next Generation Networks,” J. Netw. Syst. Manage., vol. 16, pp. 14-45, 2008.

[177] J.L. Rougier, “State of the art of Linux projects for core network emulation“,EURO NGI, Design and Engineering of the Next Generation Internet, towards convergent multiservice networks, May 2010. [Online]. Available: eurongi.enst.fr/archive/127/JRA421.pdf

[178] J. Ruiz, A. Vallejo, J. Abella, “IPv6 Conformance and Interoperability Testing”, IEEE Symposium on Computers and Communications (ISCC'05), 10th ISCC, pp. 83-88, June 2005.

[179] R. Sahita, Ed., S. Hahn, K. Chan, K. McCloghrie, “Framework Policy Information Base”, IETF RFC-3318, March 2003.

[180] S Salsano, E Sangregorio, M Listanti, “COPS DRA: a protocol for dynamic Diffserv Resource Allocation.”, Joint Planet-IP NEBULA workshop, 2002.

[181] V. Sarangan, J. Chen: "Comparative study of protocols for dynamic service negotiation in the next-generation Internet", IEEE Communications Magazine, March 2006, Vol. 44, N. 3, pp. 151-156.

[182] J.M. Selga, A. Zaballos, J. Abella, G. Corral, "Model for polling in noisy multihop systems with application to PLC and AMR," Computers and Communications, 2008. ISCC 2008. IEEE Symposium on, vol., no., pp.664-669, 6-9 July 2008.

[183] SPiDCOM, “Powerline Communication by SPiDCOM,” May 2010 [Online]. Available: www.spidcom.com/techno1.html

[184] Spirent Communications, Spirent AX/4000, May 2010, [Online]. Available: www.spirentcom.com

[185] M. Söderqvist, “Search Heuristics for Load Balancing in IP Networks”, SICS. Technical Report T2005:04, March 2005.

[186] S. Sohail, S. Jha, “The Survey of Bandwidth Broker”, Technical Report UNSW CSE TR 0206, School of Computer Science and Engineering, May 2002.

[187] J. Song; M. Y. Chang; S. S. Lee; J. Joung, "Overview of ITU-T NGN QoS Control," Communications Magazine, IEEE, vol.45, no.9, pp.116-123, September 2007.

[188] TAHI project, Conformance and Interoperability Tests, May 2010, [Online]. Available: www.tahi.org

[189] Taiwan Exec Yuan National Inforamtion and Communication Infrastrue (NICI), NICI IPv6 Standard and Testing Division Laboratory, May 2010, [Online]. Available: http://interop.IPv6.org.tw

[190] T. Tamura, M. Sasaki, K, Sato, M. Motono, "Activities of IPv6 Ready Logo Program", NTT Technical Review, Vol. 7 No. 12, Dec. 2009.

[191] Telecommunications Technology Association, TTA Laboratory, May 2010, [Online]. Available: www.tta.or.kr

[192] TEQUILA project, “SrNP: Service Negotiation Protocol”, May 2010, [Online]. Available: www.ist-tequila.org/deliverables

[193] A Terzis, L Wang, J Ogawa, L Zhang, “A two-tier resource management model for the Internet”. IEEE Global Telecomunications Conference (GLOBECOM'99), 42nd GLOBECOM, Vol 3, pp: 1779-1791, December 1999.

[194] S. Thomson, T. Narten, "IPv6 Stateless Address Autoconfiguration", IETF RFC-4862, September 2007.

[195] J. Tian, L. Zhongcheng, “The next generation Internet protocol and its test”, IEEE International Conference on Communications (ICC’01), 36th ICC, Vol. 1, pp. 210-215, June 2001.

[196] Université Catholique de Louvain and the Université de Liège, TOTEM project, May 2010, [Online]. Available: http://totem.info.ucl.ac.be

[197] University of Kansas, Information and Technology Telecommunication Centre, “bandwidth broker Implementation”, [Online]. Avaliable: www.ittc.ukans.edu/-kdraoi/BB

[198] University of New Hampshire, InterOperability Laboratory, May 2010, [Online]. Available: www.iol.unh.edu

[199] University of New Hampshire, Moonv6 Project, May 2010, [Online]. Available: www.moonv6.org

[200] University of Tokyo, USAGI (UniverSAl playGround for IPv6) Project, May 2010, [Online]. Disponible www.linux-IPv6.org

[201] U.S. Department of Energy, “The Smart Grid: An Introduction,” May 2010, [Online]. Available: www.oe.energy.gov

[202] A. Vallejo, A. Zaballos, J. Abella, J.M. Selga, C. Duz, “Performance of a policy-based management system in IPv6 networks using COPS-PR”, International Conference on Networking (ICN'07), 6th ICN 2007, April 2007.

[203] A. Vallejo, A. Zaballos, J. Abella, G. Villegas, J.M. Selga, “Evaluation of a Policy-Based QoS Management Architecture over an IPv6 DiffServ testbed”, IEEE International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TridentCom '07), 3rd Tridentcom 2007, May 2007.

[204] A. Vallejo, J. Ruiz, A. Zaballos, J. Abella, J.M. Selga, “State of the art of IPv6 conformance and interoperability testing”, Communications Magazine, IEEE, vol.45, no.10, pp.140-146, October 2007.

[205] A. Vallejo, A. Zaballos, X. Canaleta, J. Dalmau, "Estudio de la gestión de la QoS extremo-extremo en la arquitectura ITU-T IMS/NGN", VII Jornadas de Ingeniería Telemática, JITEL 2008, September 2008.

[206] A. Vallejo, A. Zaballos, X. Canaleta, J. Dalmau, "End-to-End QoS Management Proposal for the ITU-T IMS/NGN Architecture", Software, Telecommunications and Computer Networks, 2008. SoftCOM 2008. 16th International Conference on, vol., no., pp.147-151, September 2008.

[207] A. Vallejo, A. Zaballos, D. Vernet, D. Cutiller, J. Dalmau, "Implementation of Traffic Engineering in NGNs using Hybrid Genetic Algorithms", Systems and Networks Communications, 2008. ICSNC '08. 3rd International Conference on, vol., no., pp.262-267, October 2008.

[208] A. Vallejo, A. Zaballos, D. Vernet, A. Orriols-Puig, J. Dalmau, "A Traffic Engineering proposal for ITU-T NGNs using Hybrid Genetic Algorithms", Advances in Networks and Services, International Journal of, vol 2, no 1, February 2009.

[209] A. Vallejo, "A new approach for Traffic Engineering for MPLS DS-TE using Hybrid Genetic Algorithms", Technical report, May 2009.

[210] A. Vallejo, A. Zaballos, A. Campos, J. Dalmau, "Optimizing the usage of COPS protocol in ITU-T NGN architecture", AICT, May 2010. In press.

[211] A. Vallejo, A. Zaballos, J.M. Selga, J. Dalmau, "Next Generation QoS control architectures for Broadband Power Lines Networks", Submitted to IEEE Transactions on Power Delivery Journal.

[212] A. Veà, " Historia, Sociedad, Tecnología y Crecimiento de la Red. Una aproximación divulgativa a la realidad más desconocida de Internet", PhD thesis, URL, 2002.

[213] S. Waldbusser, J. Saperia, T. Hongal, “Policy Based Management MIB”, IETF RFC-4011, March 2005.

[214] X. Wang, H. Schulzrinne, “RNAP: a Resource Negotiation and Pricing Protocol,” Proc. Int’l. Wksp. Network and Op. Sys. Support for Digital Audio and Video (NOSSDAV), Basking Ridge, NJ, pp. 77–93, jun. 1999.

[215] Y. Wang, Z. Wang, L. Zhang, “Internet Traffic Engineering without Full Mesh Overlaying”, IEEE Global Communications Conference (GLOBECOM’01), 20th GLOBECOM, Vol. 1, pp. 567-571, April 2001.

[216] WIDE project, TAHI Project, KAME Project i USAGI Project, May 2010, [Online]. Available: www.wide.ad.jp

[217] X. Xiao, L. Ni, “Internet QoS: A Big Picture”, IEEE Network, pp. 8-18, March/April 1999.

[218] XORP Team, “eXtensible Open Router Platform”, May 2010, [Online]. Available: www.xorp.org

[219] H. Xu, et al., "A Modification to COPS to Improve Implementation of Push Mode", draft-xu-cops-push-00.txt, feb. 2007.

[220] Z. Yujun, L. Zhongcheng, “A new formal test suite specification language for IPv6 conformance testing”, International Conference on Communication Technology (ICCT ‘03), 1st ICCT, Vol. 1, pp. 174-177, April 2003.

[221] A. Zaballos, A. Vallejo, G. Ravera, J.M. Selga, "Simulation and modeling of the coexistence of polling and contention in PLC based AMR systems", Power Line Communications and Its Applications, 2007. ISPLC 2007. IEEE International Symposium on, vol., no., pp.110-115, March 2007.

[222] A. Zaballos, A. Vallejo, J.M. Selga, X. Canaleta, "Aplicación de AGs en el encaminamiento con QoS en redes USN Access Networks", VII Jornadas de Ingeniería Telemática, JITEL 2008, September 2008.

[223] A. Zaballos, A. Vallejo, M. Majoral, J.M. Selga, "Survey and Performance Comparison of AMR over PLC standards", Power Delivery, IEEE Transactions on, vol.24, no.2, pp.604-613, April 2009.

[224] A. Zaballos, A. Vallejo, J.J. Jimenez, J.M. Selga, "QoS Broker based architecture design for the PLC access network", Power Line Communications and Its Applications, 2009. ISPLC 2009. IEEE International Symposium on, vol., no., pp.205-210, April 2009.

[225] A. Zaballos, A. Vallejo, P. Terradellas, J.M. Selga, "A genetic-based QoS aware routing for ubiquitous sensor networks", Wireless and Mobile Computing, Networking and Communications, 2009. WIMOB 2009. IEEE International Conference on, vol., no., pp.129-134, October 2009.

[226] A. Zaballos, A. Vallejo, J.J. Jimenez, J.M. Selga, "QoS Broker based management for heterogeneous Smart Electricity Networks", IEEE Symposium on Computers and Communications (ISCC'10), June 2010. In press

[227] A. Zaballos, A. Vallejo, M. Majoral, J.M. Selga, "A new paradigm for smart grids communications architecture", Submitted to IEEE Transactions on Industrial Electronics, 2010.

[228] S. Zeadally, L. Raicu, “Evaluating IPv6 on Windows and Solaris”, IEEE Internet Computing, Vol.7, No.3, pp. 51-57, May/June 2003.

[229] G. Zorn, et al., “Diameter Quality of Service Application”, draft-ietf-dime-diameter-qos-05.txt, jul. 2007.

[230] 3rd Generation Partnership Project, ETSI, May 2010, [Online]. Available: www.3gpp.org

[231] 3rd Generation Partnership Project, TS 23.207, “End-to-end QoS concept and architecture”, v. 7.0.0., Sophia-Antipolis, 2007.

[232] 6POWER project, May 2010 [Online]. Available: www.6power.org

Appendix A: NOTATION

ABR: Available Bit Rate.

AC: Access Category.

AF: Assured Forwarding.

AI: Artificial Intelligence.

AIFS: Arbitration InterFrame Space.

AMR: Automatic Meter Reading.

ARPU: Average Revenue Per User.

ATM: Asynchronous Transfer Mode.

BA: Behavior Aggregate.

BB: Bandwidth Broker.

BE: Best Effort.

BPL: Broadband Power Line networks.

CAC: Connection Admission Control.

CAPEX: Capital Ependitures.

CAR: Committed Access Rate.

CBQ: Class Based Queuing.

CBR: Constant Bit Rate.

CFP: Contention Free Period.

CIDR: Classless Inter-Domain Routing.

COPS: Common Open Policy Server.

COPS-PR: Common Open Policy Server-Provisioning.

COPS-SLS: Common Open Policy Server - Service Level Specification.

CoS: Class of Service.

CP: Contention Period.

CPE: Customer Premises Equipment.

CSCF: Call State Control Function.

CSMA/CA: Channel Sense Multiple Access / Collision Avoidance.

CW: Contention Window.

DiffServ: Differentiated Services.

DHCP: Dynamic Host Configuration Protocol.

DSNP: Dynamic Service Negotiation Protocol.

ECMP: Equal Cost Multiple Path.

EDCA: (HFC) Enhanced Distributed Channel Access.

EF: Expedited Forwarding.

ETSI: European Telecommunications Standards Institute.

FDR: Frequency Division Repeater.

GA: Genetic Algorithm.

GNU: GNU's Not Unix.

GRED: Generic Random Early Detection.

HC: Hybrid Controller.

HCCA: HCF Controlled Channel Access.

HE: Head End.

HFC: Hybrid Coordination Function.

HGA: Hybrid Genetic Algorithm. *also* Memetic algorithm

IAB: Internet Architecture Board.

IANA: Internet Assigned Numbers Authority.

ICANN: Internet Corporation for Assigned Names and Numbers.

I-CSCF: Interrogating-CSCF.

IED: Intelligent Electronic Device.

IEEE: Institute of Electrical and Electronics Engineers.

IESG: Internet Engineering Steering Group.

IETF: Internet Engineering Task Force.

IKE: Internet Key Exchange.

IMS: IP Multimedia Subsystem.

IntServ: Integrated Services.

IPng: Internet Protocol next generation.

IPSec: Internet Protocol Security.

IPTV: Internet protocol television.

IPv6: Internet Protocol version 6.

IPv6TF: IPv6 Task Force.

IRTF: Internet Research Task Force.

ISOC: Internet Society.

IS-IS: Intermediate System-to-Intermediate System.

ITU-T: International Union for Telecommunications, Telecommunications Sector.

JATE: Japan Approvals Institute for Telecommunications Equipment.

LAN: Local Area Network.

LP: Linear Programming.

LSP: Label Switched Path.

LV: Low-Voltage (power network).

MAC: Medium Access Control.

MGC: Media Gateway Controller.

MGw: Media Gateway.

MIB: Management Information Base.

MIPv6: Mobile IPv6.

MLD: Multicast Listener Discovery.

MPLS: Multi Protocol Label Switching.

MPLS DS-TE: DiffServ Aware MPLS-TE.

MPLS-TE: Multi Protocol Label Switching - Traffic Engineering.

MV: Medium-Voltage (power network).

NAv6TF: North American IPv6 Task Force.

NEMO: Network Mobility.

NGA: Next Generation Access networks.

NGN: Next Generation Network.

NGN-GSI: (ITU-T) Next Generation Networks Global Standards Initiative.

NIST: National Institute of Standards and Technology.

NMS: Network Management System.

NREN: National Research and Education Network.

NSIS: Next Steps in Signaling.

NTU: Network Termination Unit.

NTRD: Network Topology and Resource Database.

NUT: Node Under Test.

OPEX: Operating Expenditures.

OSPF: Open Shortest Path First.

OSPFWS: Open Shortest Path First weight setting (problem)

OVLAN: OPERA project Virtual Local Area Network.

PBNM: Policy-Based Network Management.

PDU: Protocol Data Unit.

PDP: Policy Decision Point.

PEP: Policy Enforcement Point.

PD-FE: Policy Decision Functional Entity.

PE-FE: Policy Enforcement Functional Entity.

PHB: Per-Hop Behavior.

PIB: Policy Information Base.

PLC: Power Line Communications.

PMTUD: Path MTU Discovery.

PRC: Provisioning Class.

PSTN: Public Switched Telephone Network.

PTCC: Protocol and Testing Competence Centre.

P-CSCF: Proxy CSCF.

QoS: Quality of Service.

QoSB: QoS Broker.

QoS-GSLP: QoS Generic Signaling Layer Protocol.

QoS-NSLP: QoS-NSIS Signaling Layer Protocol.

RAA: Resource Allocation Answer.

RACF: Resource and Admission Control Functions.

RACS: Resource and Admission Control Sub-system.

RAR: Resource Allocation Request.

RED: Random Early Detection.

RFC: Request for Comment.

RNAP: Resource Negotiation and Pricing Protocol.

RP: Repeater.

RSVP: Resource ReSerVation Protocol.

RWS: Roulette Wheel Selection.

SCCF: Service and Call Control Functions.

SDP: Session Description Protocol.

SIBBS: Simple Interdomain Bandwidth Broker Specification.

SIBBSv6: Simple Interdomain Bandwidth Broker Specification for IPv6.

SIP: Session Initiation Protocol.

SLA: Service Level Agreement.

SLS: Service Level Specification.

SNMP: Simple Network Management Protocol.

SrNP: Service Negotiation Protocol.

S-CSCF: Serving CSCF.

TCP: Transmission Control Protocol.

TCS: Traffic Conditioning Specification.

TDMA: Time Division Multiple Access.

TDR: Time Division Repeater.

TISPAN: (ETSI) Telecommunications and Internet converged Services and Protocols for Advanced Networking.

TN: Tester Node.

ToIP: Telephony over IP.

ToS: Type of Service.

TR: Tester Router.

TRC-FE: Transport Resource Control - Functional Entities.

TRE-FE: Transport Resource Enforcement - Functional Entities.

TTCN-3: Testing and Test Control Notation version 3.

TXOP: Transmission Opportunity.

UDP: User Datagram Protocol.

UMTS: Universal Mobile Telephone System.

UNH-IOL: University of New Hampshire - InterOperability Laboratory.

UP: User Priorities.

URL (1): Uniform Resource Locator.

URL (2): Univeristat Ramon Llull.

UNSW: University New South Walles

USG: USA Government.

VBR: Variable Bit Rate.

VoIP: Voice over IP.

v6PC: IPv6 Promotion Council.

WAN: Wide Area Network.

WDM: Wavelength Division Multiplexing.

WFQ: Weighted Fair Queuing.

WG: Working Group.

WiMAX: Worldwide Interoperability for Microwave Access.

WRED: Weighted Random Early Detection.

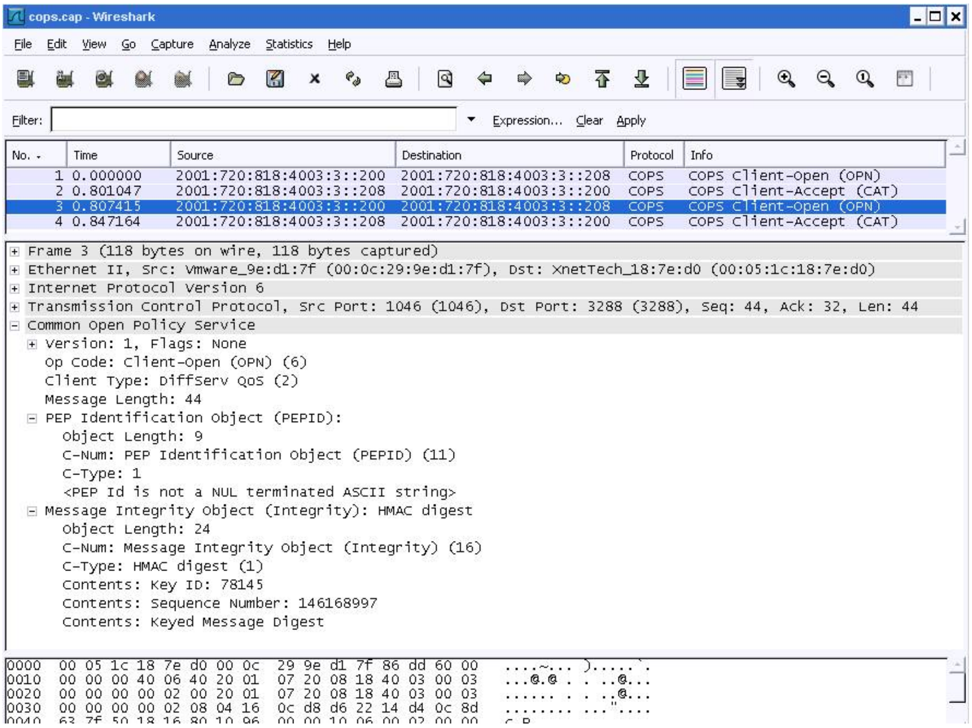
3GPP: 3rd Generation Partnership Project.

Appendix B: Chapter 3

This appendix provides some detailed information about the BBv6. It shows some sniffer captures of the COPS-PR protocol, it details the JAVA classes of the BBv6 and show the CLI administrator menu.

B.1. COPS-PR captures

The following captures have been captured with Wireshark application. This software supports the COPS-PR protocol.

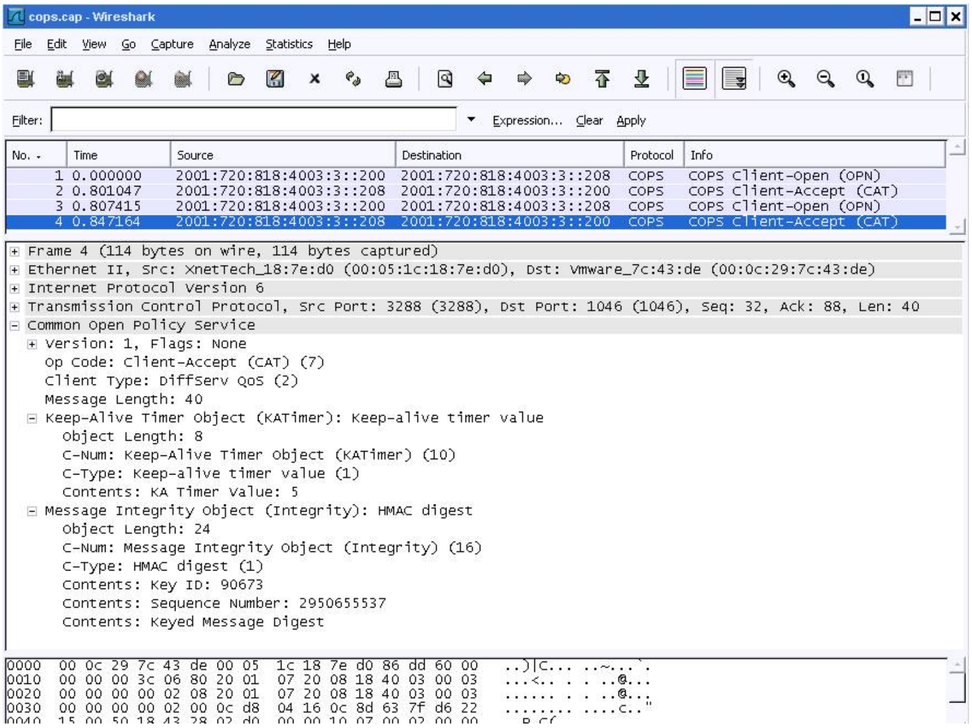

cops.cap - Wireshark
File Edit View Go Capture Analyze Statistics Help
Filter:
Expression... Clear Apply

```
No. -  Time       Source                      Destination                 Protocol  Info
   1 0.000000   2001:720:818:4003:3::200   2001:720:818:4003:3::208   COPS      COPS Client-Open (OPN)
   2 0.801047   2001:720:818:4003:3::208   2001:720:818:4003:3::200   COPS      COPS Client-Accept (CAT)
   3 0.807415   2001:720:818:4003:3::200   2001:720:818:4003:3::208   COPS      COPS Client-Open (OPN)
   4 0.847164   2001:720:818:4003:3::208   2001:720:818:4003:3::200   COPS      COPS Client-Accept (CAT)
```

```
+ Frame 3 (118 bytes on wire, 118 bytes captured)
+ Ethernet II, Src: Vmware_9e:d1:7f (00:0c:29:9e:d1:7f), Dst: XnetTech_18:7e:d0 (00:05:1c:18:7e:d0)
+ Internet Protocol Version 6
+ Transmission Control Protocol, Src Port: 1046 (1046), Dst Port: 3288 (3288), Seq: 44, Ack: 32, Len: 44
- Common Open Policy Service
  + Version: 1, Flags: None
    Op Code: Client-Open (OPN) (6)
    Client Type: DiffServ QoS (2)
    Message Length: 44
  - PEP Identification Object (PEPID):
      Object Length: 9
      C-Num: PEP Identification Object (PEPID) (11)
      C-Type: 1
      <PEP Id is not a NUL terminated ASCII string>
  - Message Integrity Object (Integrity): HMAC digest
      Object Length: 24
      C-Num: Message Integrity Object (Integrity) (16)
      C-Type: HMAC digest (1)
      Contents: Key ID: 78145
      Contents: Sequence Number: 146168997
      Contents: Keyed Message Digest
```

```
0000  00 05 1c 18 7e d0 00 0c  29 9e d1 7f 86 dd 60 00   ....~... ).....`.
0010  00 00 00 40 06 40 20 01  07 20 08 18 40 03 00 03   ...@.@ . . ..@...
0020  00 00 00 00 02 00 20 01  07 20 08 18 40 03 00 03   ...... . . ..@...
0030  00 00 00 00 02 08 04 16  0c d8 d6 22 14 d4 0c 8d   ........ ..."....
0040  63 7f 50 18 16 80 10 96  00 00 10 06 00 02 00 00   c.P.
```

Figure 87: Client-Open message of the COPS-PR protocol captured in the testbed


cops.cap - Wireshark
File Edit View Go Capture Analyze Statistics Help
Filter:
Expression... Clear Apply

```
No. -  Time       Source                      Destination                 Protocol  Info
   1 0.000000   2001:720:818:4003:3::200   2001:720:818:4003:3::208   COPS      COPS Client-Open (OPN)
   2 0.801047   2001:720:818:4003:3::208   2001:720:818:4003:3::200   COPS      COPS Client-Accept (CAT)
   3 0.807415   2001:720:818:4003:3::200   2001:720:818:4003:3::208   COPS      COPS Client-Open (OPN)
   4 0.847164   2001:720:818:4003:3::208   2001:720:818:4003:3::200   COPS      COPS Client-Accept (CAT)
```

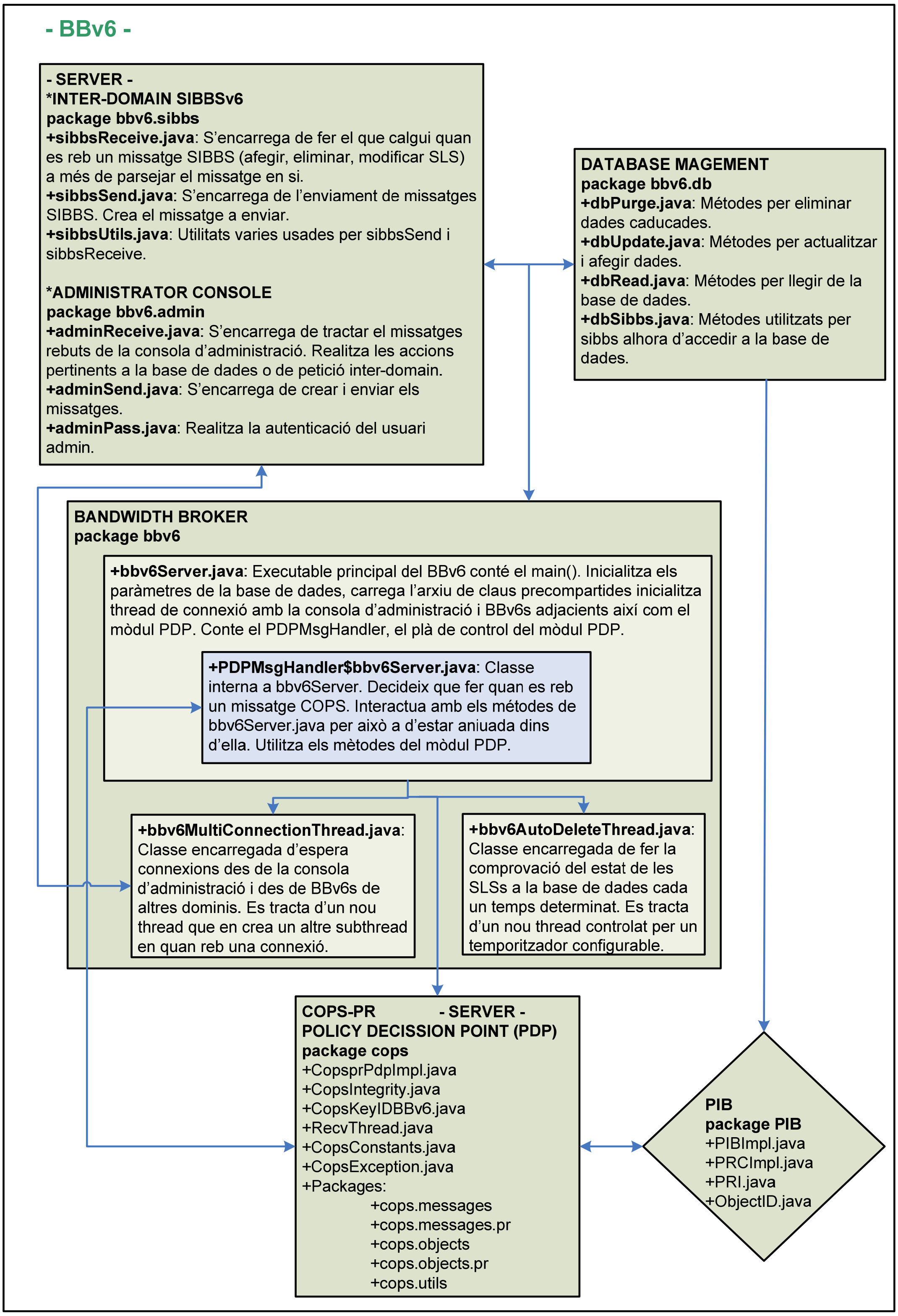
```
+ Frame 4 (114 bytes on wire, 114 bytes captured)
+ Ethernet II, Src: XnetTech_18:7e:d0 (00:05:1c:18:7e:d0), Dst: Vmware_7c:43:de (00:0c:29:7c:43:de)
+ Internet Protocol Version 6
+ Transmission Control Protocol, Src Port: 3288 (3288), Dst Port: 1046 (1046), Seq: 32, Ack: 88, Len: 40
- Common Open Policy Service
  + Version: 1, Flags: None
    Op Code: Client-Accept (CAT) (7)
    Client Type: DiffServ QoS (2)
    Message Length: 40
  - Keep-Alive Timer Object (KATimer): Keep-alive timer value
      Object Length: 8
      C-Num: Keep-Alive Timer Object (KATimer) (10)
      C-Type: Keep-alive timer value (1)
      Contents: KA Timer Value: 5
  - Message Integrity Object (Integrity): HMAC digest
      Object Length: 24
      C-Num: Message Integrity Object (Integrity) (16)
      C-Type: HMAC digest (1)
      Contents: Key ID: 90673
      Contents: Sequence Number: 2950655537
      Contents: Keyed Message Digest
```

```
0000  00 0c 29 7c 43 de 00 05  1c 18 7e d0 86 dd 60 00   ..)|C... ..~...`.
0010  00 00 00 3c 06 80 20 01  07 20 08 18 40 03 00 03   ...<.. . . ..@...
0020  00 00 00 00 02 08 20 01  07 20 08 18 40 03 00 03   ...... . . ..@...
0030  00 00 00 00 02 00 0c d8  04 16 0c 8d 63 7f d6 22   ........ ....c.."
0040  15 00 50 18 43 28 02 d0  00 00 10 07 00 02 00 00   ..P.C(.
```

Figure 88: Client-Accept message of the COPS-PR protocol captured in the testbed

B.2. JAVA classes

The JAVA classes for the six packages in the BBv6 server are listed. The colors code of this section follows the colors of the blocks of Figure 27.



- BBv6 -

- SERVER -
***INTER-DOMAIN SIBBSv6**
package bbv6.sibbs
+sibbsReceive.java: S'encarrega de fer el que calgui quan es reb un missatge SIBBS (afegir, eliminar, modificar SLS) a més de parsejar el missatge en si.
+sibbsSend.java: S'encarrega de l'enviament de missatges SIBBS. Crea el missatge a enviar.
+sibbsUtils.java: Utilitats varies usades per sibbsSend i sibbsReceive.

***ADMINISTRATOR CONSOLE**
package bbv6.admin
+adminReceive.java: S'encarrega de tractar el missatges rebuts de la consola d'administració. Realitza les accions pertinents a la base de dades o de petició inter-domain.
+adminSend.java: S'encarrega de crear i enviar els missatges.
+adminPass.java: Realitza la autenticació del usuari admin.

DATABASE MAGEMENT
package bbv6.db
+dbPurge.java: Métodes per eliminar dades caducades.
+dbUpdate.java: Métodes per actualitzar i afegir dades.
+dbRead.java: Métodes per llegir de la base de dades.
+dbSibbs.java: Métodes utilitzats per sibbs alhora d'accedir a la base de dades.

BANDWIDTH BROKER
package bbv6

+bbv6Server.java: Executable principal del BBv6 conté el main(). Inicialitza els paràmetres de la base de dades, carrega l'arxiu de claus precompartides inicialitza thread de connexió amb la consola d'administració i BBv6s adjacients així com el mòdul PDP. Conte el PDPMsgHandler, el plà de control del mòdul PDP.

+PDPMsgHandler$bbv6Server.java: Classe interna a bbv6Server. Decideix que fer quan es reb un missatge COPS. Interactua amb els métodes de bbv6Server.java per això a d'estar aniuada dins d'ella. Utilitza els mètodes del mòdul PDP.

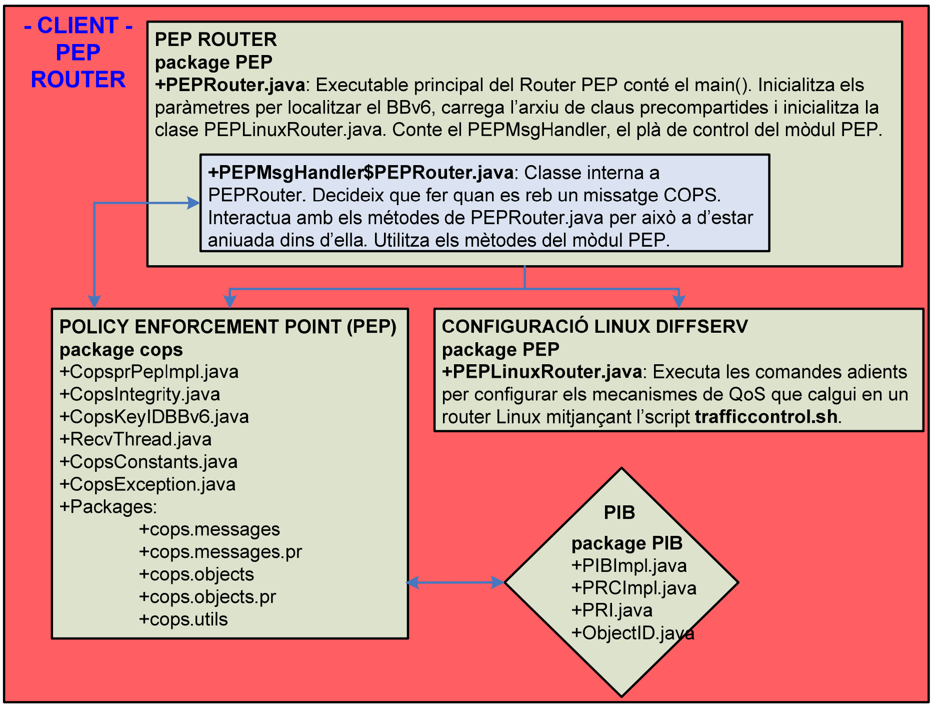
+bbv6MultiConnectionThread.java: Classe encarregada d'espera connexions des de la consola d'administració i des de BBv6s de altres dominis. Es tracta d'un nou thread que en crea un altre subthread en quan reb una connexió.

+bbv6AutoDeleteThread.java: Classe encarregada de fer la comprovació del estat de les SLSs a la base de dades cada un temps determinat. Es tracta d'un nou thread controlat per un temporitzador configurable.

COPS-PR - SERVER -
POLICY DECISSION POINT (PDP)
package cops
+CopsprPdpImpl.java
+CopsIntegrity.java
+CopsKeyIDBBv6.java
+RecvThread.java
+CopsConstants.java
+CopsException.java
+Packages:
+cops.messages
+cops.messages.pr
+cops.objects
+cops.objects.pr
+cops.utils

PIB
package PIB
+PIBImpl.java
+PRCImpl.java
+PRI.java
+ObjectID.java

The JAVA classes for the three packages in the PEP router are listed:



- CLIENT -
PEP
ROUTER

PEP ROUTER
package PEP
+PEPRouter.java: Executable principal del Router PEP conté el main(). Inicialitza els paràmetres per localitzar el BBv6, carrega l'arxiu de claus precompartides i inicialitza la clase PEPLinuxRouter.java. Conte el PEPMsgHandler, el plà de control del mòdul PEP.

+PEPMsgHandler$PEPRouter.java: Classe interna a PEPRouter. Decideix que fer quan es reb un missatge COPS. Interactua amb els métodes de PEPRouter.java per això a d'estar aniuada dins d'ella. Utilitza els mètodes del mòdul PEP.

POLICY ENFORCEMENT POINT (PEP)
package cops
+CopsprPepImpl.java
+CopsIntegrity.java
+CopsKeyIDBBv6.java
+RecvThread.java
+CopsConstants.java
+CopsException.java
+Packages:
+cops.messages
+cops.messages.pr
+cops.objects
+cops.objects.pr
+cops.utils

CONFIGURACIÓ LINUX DIFFSERV
package PEP
+PEPLinuxRouter.java: Executa les comandes adients per configurar els mecanismes de QoS que calgui en un router Linux mitjançant l'script **trafficcontrol.sh**.

PIB
package PIB
+PIBImpl.java
+PRCImpl.java
+PRI.java
+ObjectID.java

The summary of the JAVA classes for the COPS-PR communication protocol are listed:

COPS-PR - SERVER -
POLICY DECISSION POINT (PDP)
package cops
+CopsprPdpImpl.java
+CopsIntegrity.java
+CopsKeyIDBBv6.java
+RecvThread.java
+CopsConstants.java
+CopsException.java
+Packages:
+cops.messages
+cops.messages.pr
+cops.objects
+cops.objects.pr
+cops.utils

COPS-PR - CLIENT -
POLICY ENFORCEMENT POINT (PEP)
package cops
+CopsprPepImpl.java
+CopsIntegrity.java
+CopsKeyIDBBv6.java
+RecvThread.java
+CopsConstants.java
+CopsException.java
+Packages:
+cops.messages
+cops.messages.pr
+cops.objects
+cops.objects.pr
+cops.utils

The JAVA classes for the administrator console are listed:

-CLIENT-
ADMINISTRATOR CONSOLE

INTERFÍCIE ADMINISTRADOR
package console
+consoleAdmin.java: Executable principal conté el main(). Realitza la connexió i autenticació amb el BBv6. Pinta el menú per pantalla i s'encarrega de cridar als métodes corresponents a cada opció del menú disponible. Pot afegir, modificar i eliminar tant SLAs com SLSs.
+consoleSend.java: S'encarrega de preparar els missatges a enviar per cada opció.
+consoleReceive.java: S'encarrega de parsejar el missatge rebut i mostrar la informació de manera entenedora per pantalla.

B.3. Console Administration Menu

The CLI menu for the remote administration of the BBv6 is show next.

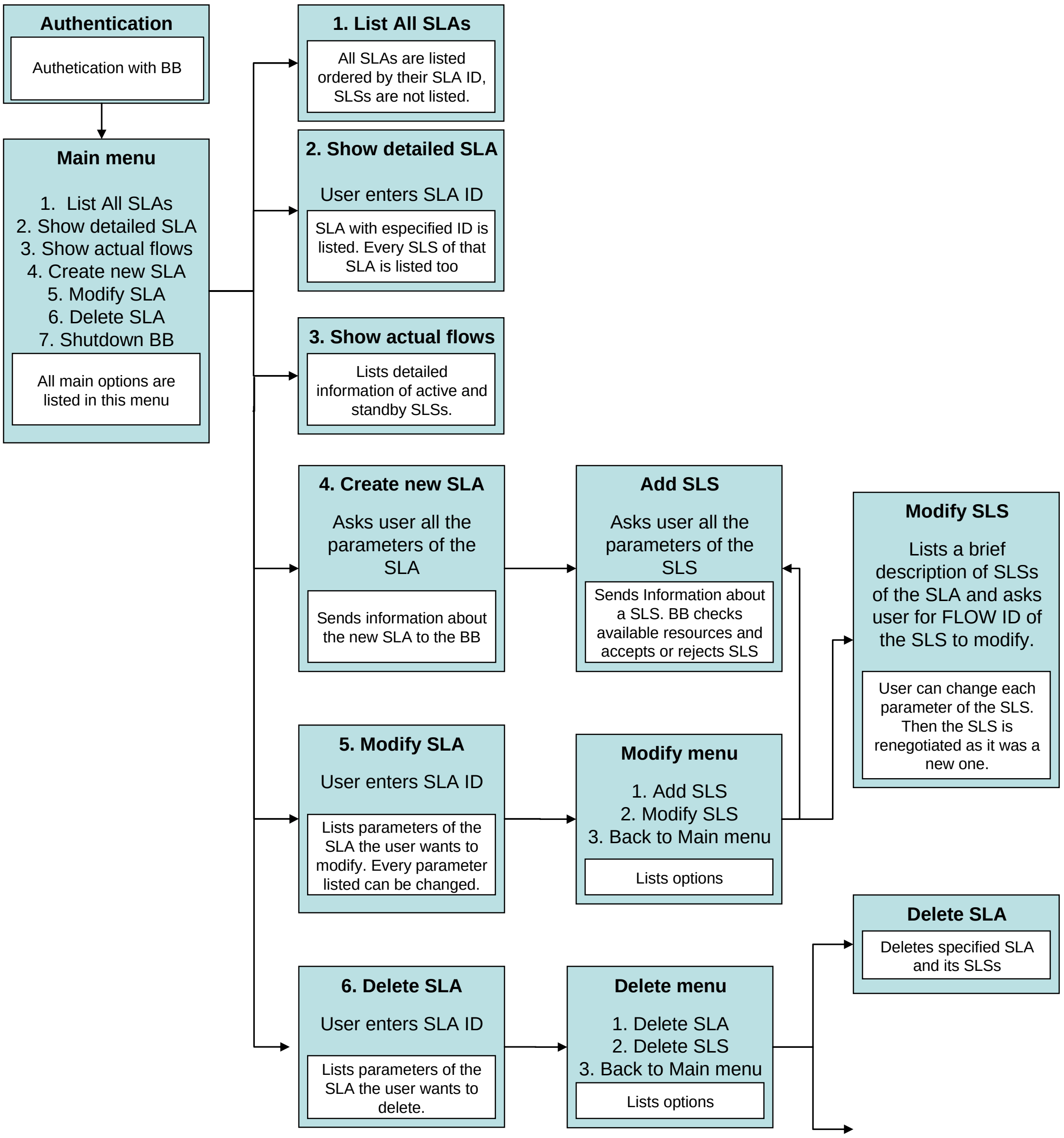

Authentication
Authetication with BB
Main menu
1. List All SLAs
2. Show detailed SLA
3. Show actual flows
4. Create new SLA
5. Modify SLA
6. Delete SLA
7. Shutdown BB
All main options are listed in this menu
1. List All SLAs
All SLAs are listed ordered by their SLA ID, SLSs are not listed.
2. Show detailed SLA
User enters SLA ID
SLA with especified ID is listed. Every SLS of that SLA is listed too
3. Show actual flows
Lists detailed information of active and standby SLSs.
4. Create new SLA
Asks user all the parameters of the SLA
Sends information about the new SLA to the BB
Add SLS
Asks user all the parameters of the SLS
Sends Information about a SLS. BB checks available resources and accepts or rejects SLS
Modify SLS
Lists a brief description of SLSs of the SLA and asks user for FLOW ID of the SLS to modify.
User can change each parameter of the SLS. Then the SLS is renegotiated as it was a new one.
5. Modify SLA
User enters SLA ID
Lists parameters of the SLA the user wants to modify. Every parameter listed can be changed.
Modify menu
1. Add SLS
2. Modify SLS
3. Back to Main menu
Lists options
Delete SLA
Deletes specified SLA and its SLSs
6. Delete SLA
User enters SLA ID
Lists parameters of the SLA the user wants to delete.
Delete menu
1. Delete SLA
2. Delete SLS
3. Back to Main menu
Lists options

Figure 89: Menu structure of the BBv6's console administration

Appendix C: Chapter 6

This appendix provides some detailed information of the Chapter 6 testing. It shows the tables used to build the graphics and the first testbed implemented with Cisco Sustems routers.

C.1. Simulation results for N11

This section provides the tables used to build the graphics of HGA comparison in Chapter 6.

Average results with ECMP

	Traffic demand				
	15.000	20.000	25.000	35.000	50.000
Inverse capacity [27]	7.94	10.47	13.98	17.90	31.23
Fortz [70]	7.64	10.03	13.45	16.94	30.83
Ericsson [47]	7.56	10.05	13.40	16.63	30.91
Mulyana [146]	6.93	9.52	12.74	16.16	28.93
Buriol [17]	7.37	10.02	13.50	17.32	30.98
Riedl [175]	7.48	10.11	13.74	17.41	30.95

Maximum results with ECMP

	Traffic demand				
	15.000	20.000	25.000	35.000	50.000
Inverse capacity [27]	50.00	53.00	63.00	79.14	139.40
Fortz [70]	26.40	39.17	57.24	63.83	93.06
Ericsson [47]	30.30	34.64	58.04	64.10	96.39
Mulyana [146]	23.20	29.45	43.12	52.81	90.10
Buriol [17]	25.20	30.99	40.00	51.62	85.75
Riedl [175]	22.70	28.69	42.68	51.91	87.44

Average results without ECMP

	Traffic demand				
	15.000	20.000	25.000	35.000	50.000
Inverse capacity [27]	8.13	10.85	14.74	18.89	32.44
Fortz [70]	8.41	9.83	13.1	15.88	32.03
Ericsson [47]	7.32	9.6	12.49	15.58	30.6
Mulyana [146]	6.49	8.64	12.55	14.6	28.76
Buriol [17]	7.82	9.6	12.29	15.55	30.43
Riedl [175]	7.33	10.48	13.76	16.97	30.23

Maximum results without ECMP

	Traffic demand				
	15.000	20.000	25.000	35.000	50.000
Inverse capacity [27]	50	56.66	73.33	83.33	156.6
Fortz [70]	35.7	52.72	55.35	74.28	90
Ericsson [47]	40.3	53.13	57.04	71.42	90.2
Mulyana [146]	35.2	50	50.28	71.42	91.13
Buriol [17]	32.7	51.97	50.8	70.92	86.62
Riedl [175]	35.7	50	50.28	71.42	90.7

C.2. First testbed implemented with Cisco System routers

In order to test the routers, and the new traffic optimization application, was built a first testbed with only seven routers.

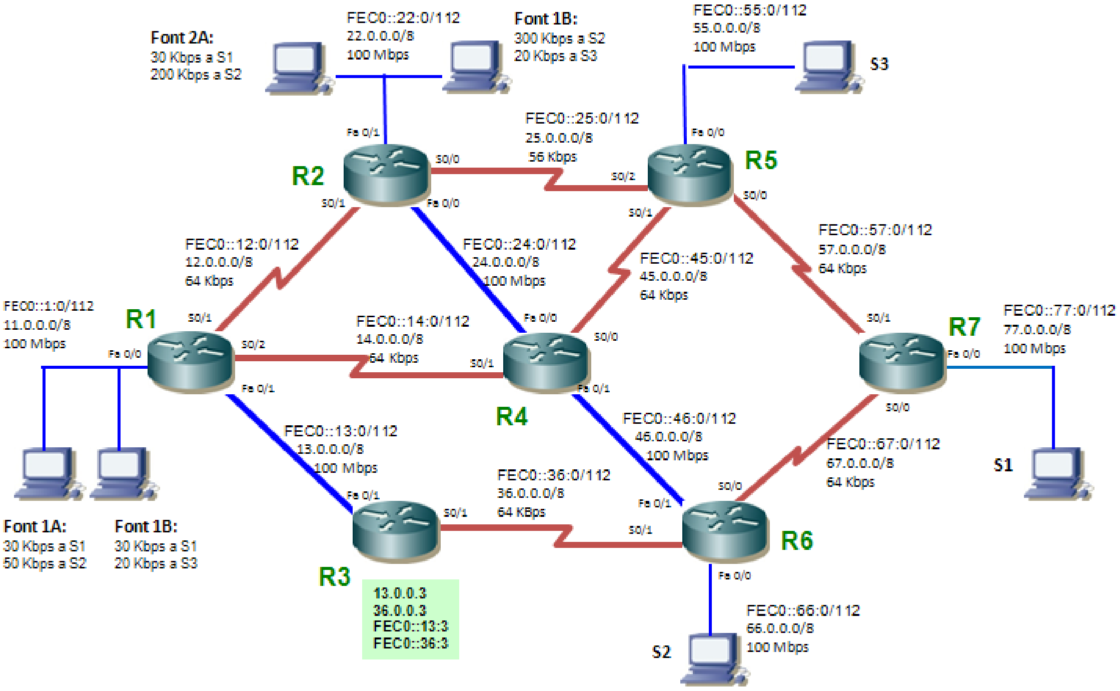


Figure 90: Fist testbed implemented with seven Cisco System routers

Appendix D: State of the art of MPLS-TE implementations

This appendix provides a state of the art study of MPLS-TE implementations mentioned in the Future work section. It is part of a technical report used as a deliverable for the ENIGMA 3 project [45].

D.1. Introduction

This document analyzes the state of the art of MPLS-TE in GNU/Linux environments. We discuss the different open source and commercials applications and daemons that have been deployed. Special attention might be paid to section 6 that describes the comercial platforms developed by different companies that are suitable to support features related with project Enigma.

Along this document we will see that open source implementations of MPLS-TE are all incomplete, will a lot of features not yet implemented, so these implementations are no suitable to use in ENIGMA III. The main cause why Open Source implementations don’t evolve quickly is that there aren’t enough developers and efforts; for instance, MPLS-Linux: that is the most important project to bring a Linux Router the capacity of act as MPLS router, has only one developer and the future for open source implementations of MPLS is not very clear.

So, the only choice left is to use commercial implementations, where most of the last RFCs and drafts are implemented. The only problem is the prize of these implementations.

D.2. MPLS-TE Testbed Linux Architecture

Core networks involve by nature very high- speed equipments that can’t be realistically emulated by common PC hardware. For instance, with the evolution of the Cisco hardware platforms, ASICs designs, and new bus systems, a lot of functionality has been delegated from the CPU to linecards, daughter cards, and custom chips. Cisco has also done a lot of development in the area of fast-switching strategies (Cisco Express Forwarding, silicon switching, fast switching), whereas almost everything in the IP stack of UNIX operating systems is done on a per-packet or per-frame basis with different per-flow characteristics.

In opposition, the control plane of core network equipments is very close to a computer, in term of both software architecture used and hardware. In this case, open source test beds offer great opportunities for various applications:

- Early protocol/architecture implementation.
- Validation of protocols, or Traffic Engineering tools in realistic environment.
- Test of scalability and performance in large scale networks: given the reduced cost of standard PCs, it can be affordable to build large scale test beds for control plane emulation.

Design of Traffic Engineering Module

The design of a traffic engineering module involves various distinct components that work together to provide constraint based routing facility. Additional link information such as the residual bandwidth or the bandwidth reserved for backup paths must be propagated through OSPF or ISIS with TE extensions. Once this information is propagated, it must be retrieved and subsequently used by a constraint based routing application. This application must make use of this additional information to calculate shortest paths that also satisfy the given constraint. Once such a route is calculated, it must be signalled so that the required resources can be reserved. This signalling of the path for resource reservation is done through signalling protocol like RSVP-TE or CR-LDP. Finally, the constraint based routing application at all the nodes should ensure that updated link state information is again propagated through OSPF-TE [169]. Thus, the constraint based routing application acts as an interface between OSPF-TE and RSVP-TE, in addition to calculating constraint based shortest paths.

Big Picture of the Architecture

The MPLS-TE control plane is responsible for (1) distribution routing information among LSRs, (2) the algorithms that these routers use to convert this information into a forwarding table that is used by the forwarding (transport/user) plane and (3) the procedures for mapping between labels and next hops (label distribution). In brief, such a control plane consists of conventional network layer routing protocols (OSPF, BGP, ISIS...) and one or more label binding mechanisms [177].

Regarding to the MPLS control plane, there are several components which should be available to support Traffic Engineering. Each component has a certain role in process of LSPs establishment, among which these ones are the most important are depicted in next figure (inspired from a tutorial from Tellium Inc on GMPLS control plane development):

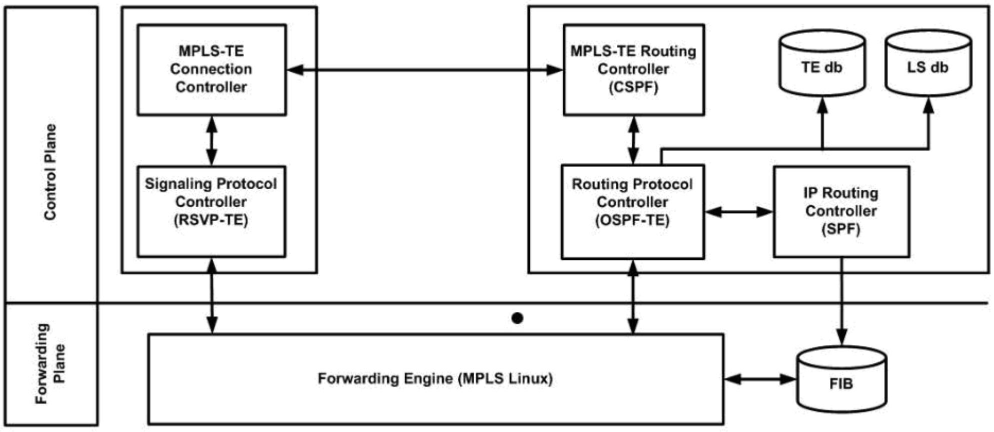

Control Plane
Forwarding Plane
MPLS-TE Connection Controller
Signaling Protocol Controller (RSVP-TE)
MPLS-TE Routing Controller (CSPF)
TE db
LS db
Routing Protocol Controller (OSPF-TE)
IP Routing Controller (SPF)
Forwarding Engine (MPLS Linux)
FIB

Figure 91: Implementation of MPLS-TE control plane on a Linux architecture

MPLS-TE Connection Controller: this component is in charge of set-up, maintenance and release of MPLS-TE connections. It keeps track of all connections (for each it is the ingress, egress or just an intermediate LSR). As an ingress (responsible of setting up a new connection), it uses the Routing Protocol Controller for path selection (Explicit Route Object ERO determination) and then of the Signalling Protocol Controller for signalling the connection.

Routing Protocol Controller: this component is in charge of discovering neighbour routers and exchanging routing information. In reality, Traffic Engineering relies on the IGP extensions (e.g. OSPF-TE) to distribute/flood link-related information about resource availability of links such as bandwidth per priority (0-7), TE specific link metric, and resource class affinity, etc. Each LSR creates and maintains a TE Link State database (TE-LSDB) that is updated by IGP flooding whenever a change occurs.

Routing Controller: this component is in charge of route computation, i.e. it implements Constrained Shortest Path First (CSPF) which aims to find the optimal path for a LSP tunnel that meets specified constraints. In CSPF, the route setup happens at the ingress router; whereas, in conventional IP Routing each router must compute the path based on its own routing information.

CSPF is originated only when (1) a new LSP establishment is requested; (2) a node/link on the current LSP has failed; (3) to re-optimize a LSP. CSPF considers the information concerning the attributes associated with traffic trunks and resources as well as the topology state that are maintained by the Routing Protocol Controller (OSPF-TE).

Signalling Protocol Controller: this component is in charge of handling the RSVP-TE signalling messages (PATH and RESV, etc.). Labels are allocated from the downstream direction.

In figure, it is describe the functionalities and the relation between the aforementioned components of a MPLS-TE ingress router. After receiving a request for establishing a LSP, module MPLS-TE Connection Controller uses MPLS-TE Routing Controller in order to calculate an optimal route that meets all requirements by executing a CSPF calculation. CSPF, basing on TE-LSDB, will compute an explicit route then forwards this ERO and its constraints to module Signalling Protocol Controller which uses signalling protocol such as RSVP-TE to distribute the labels. Once the LSP is established, the next step is to forward traffic across this LSP.

Note that for the sake of simplicity we concentrated only the main functional blocks. The management plane is also important for experimentations (i.e. for off-line Traffic Engineering methods validation) or for policy enforcement. There is also a need for such interfaces (SNMP or COPS for instance).

D.3. Open Source Routing Suites

From different approaches of open source routing platforms in this study we focused in Zebra and its later evolution (fork) Quagga. This is due to its no wonder that it has biggest feature set. Really the suite of routing protocols. It provides most of stuff needed to implement new protocol and is the more developed for the community [79].

For other interesting -commercial and no commercial- suites like XORP [218] please refer to addendum's.

Zebra

The idea for Zebra originally came from Kunihiro Ishiguro, after he realized the need for quality routing software. While working at an ISP joint venture between British Telecom and Marubeni, Ishiguro encountered venture capitalist Yoshinari Yoshikawa.Yoshinari Yoshikawa shared Ishiguro's vision for a new routing engine and they combined resources to create the world's first routing engine software. The resulting entity, known as the Zebra Project, was started in 1996.

Since its last release (version 0.95) in 2005-09-08, development for Zebra has appeared to have been stopped. A new project has emerged as the unofficial successor of Zebra: Quagga [168]. Worldwide, many Linux/BSD based BGP-routers which were using Zebra are switching to Quagga.

Zebra is distributed multi-server multi-threaded routing software and is free software distributed under GPL (GNU Public License). Traditional routing softwares (such as GateD) are made as one process program that provides all of the routing protocol functionalities as a whole. Zebra is unique in its design in which it has a process (running in background, thus a daemon) for each protocol. Zebra

uses multithread technology under multithread supported Unix kernels. Thus Zebra provides flexibility and reliability. Each module can be upgraded independently of the others, allowing for quick upgrades as well as protection from the case of a failure in one protocol affecting the entire system.

These daemons include Zebra daemon (which acts as a central arbiter) and other protocol daemons such as, a RIP daemon (RIPd), OSPF daemon (OSFPd), and a BGP daemon (BGPd).

Protocol daemons do not communicate with the kernel directly. Instead, Zebra defines its own set of protocol (known as Zebra Protocol) to handle inter-process communication between the Zebra daemon and other daemons, and then, the Zebra daemon is responsible to communicate with the Linux kernel.

As a result, Zebra mimics the client-server model in which the protocol daemons are clients and the Zebra daemon is a server that allocates and distributes services and resources (routing table information) from the kernel to its client.

Each daemon has its own routing table. Zebra daemon is responsible for kernel routing table update (service) and its redistribution between different protocol (e.g. RIPd). Other daemons are for protocol handling.

Zebra achieves modularity, extensibility, and maintainability in its architecture design. This leads to advantage as well as disadvantages.

Advantages and Disadvantages

Advantages

- Modularity: due to the multi-process nature of the Zebra software, it is easily upgraded and maintained. Each protocol can be upgraded separately, leaving the other protocols and the router online. This will save network administrators time in upgrading and maintenance.
- Speed: packet routing is carried out at a faster rate than with traditional software. Zebra software allows routers to transfer more data quicker. The need for the ability to transfer large amounts of data quickly is increasing as the internet grows and global networks form. Zebra software will meet that need.
- Reliability: in the event of failure of any of the software modules, the router can remain online and the other protocol daemons will continue to operate. The failure can then be diagnosed and corrected without taking the router offline.

Disadvantages

- Inter-process communication (Zebra) v.s. intra-process communication (GateD).
- Kernel routing table updates and redistribution would be slow.

For study zebra and its architecture in-depth we refer to [64].

Quagga

As we mentioned before Quagga is a fork of the GNU Zebra project (inactive since 2005). The Quagga tree aims to build a more involved community around Quagga than the current centralised model of GNU Zebra. It is these Zserv clients which typically implement a routing protocol and communicate routing updates to the zebra daemon. Existing Zserv clients are: ospfd (implementing OSPFv2); ripd (implementing RIP v1 and V2); ospf6d (implementing OSPFv3 - (IPv6)); ripngd (implementing RIP ng (IPv6)); bgpd (implementing BGPv4+ (including address family support for multicast and Ipv6)).

Additionally, the Quagga architecture has a rich development library to facilitate the implementation of protocol/client daemons, coherent in configuration and administrative behaviour. Currently, Quagga supports common unicast routing protocols. Multicast routing pro-tocols such as BGMP, PIM-SM, PIM-DM may be supported in Quagga 2.0. MPLS support is going on. In the future, TCP/IP filtering control, QoS control, diffserv configuration will be added to Quagga. Quagga project s final goal is making a productive, quality, free TCP/IP routing software.

D.4. Open Source Daemons Implementations

Now we will explain some of the open source daemons implementations that it’s related to MPLS and traffic engineering.

OSPFv2 daemon with TE extensions

Support for TE and OSPF-API Ospfd implements OSPF version 2 (RFC 2328). It provides the basic protocol functions such as neighbor discovery, initial link-state database exchange, propagating topology changes to neighbors, building internal link-state database (LSDB) and path computation from LSDB.

To support Traffic Engineering extensions, ospfd was added modules opaque LSAs (RFC 2370) and TE (RFC 3630 Traffic Engineering (TE) Extensions to OSPF Version 2). A new set of TLV was proposed, among which we may cite the Link Type TLVs (type-length-value) which identifies the type of link and its Traffic Engineering parameters such as unreserved bandwidth per priority (0-7), TE link metric and color, etc. Note however that TE information are exchanged as opaque LSA (as specific in IETF standards). When enabled, TE links (Link States providing TE information) are actually announced within the routing area and stored in a specific TE database (in fact the opaque-lsa database of OSPFD). However, standard IP routing computation does not take into account these information, by nature (they are opaque link states). These information can however been used for MPLS-TE LSP route selection process, which is not provided in the Quagga suite. However, OSPF-API, described below, offers an excellent tool in order to access the ospfd database (including opaque lsa) and thus to build such routing algorithms.

Accessing OSPFD Database: OSPF-API [161]. Since ospfd already keeps track of the network topology, it s important to obtain this information by accessing its LSDB. To gain access to the ospfd, there is a need for an API that provides external applications with the following functionality: retrieval of the full or partial link-state database of the OSPF daemon. This allows applications to obtain an exact copy of the LSDB including router LSAs, network LSAs and so on. Whenever a new LSA arrives at the OSPF daemon, the API module immediately informs the application by sending a message. This way, the application is always synchronized with the LSDB of the OSPF daemon, Origination of own opaque LSAs (of type 9, 10, or 11) which are then distributed transparently to other routers within the flooding scope and received by other applications through the OSPF API. The information contained in opaque LSAs is transparent for the routing process but it can be processed by other modules such as traffic engineering (e.g., MPLS-TE).

The OSPF daemon is extended with opaque LSA capabilities and an API for external applications. The OSPF core module executes the OSPF protocol by discovering neighbors and exchanging neighbor state. The opaque module provides functions to exchange opaque LSAs between routers. Opaque LSAs can be generated by several modules such as the MPLS-TE module or the API server module [176].

These modules then invoke the opaque module to flood their data to neighbours within the flooding scope.

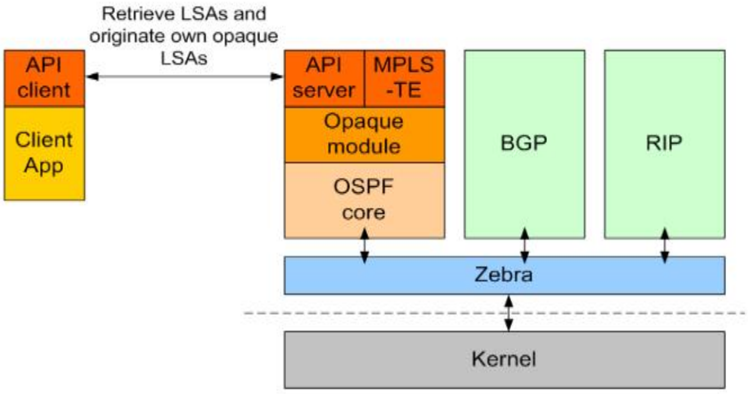


Figure 92: Ospf-api architecture

RSVP-TE Daemons

There are two major implementations of RSVP signalling daemons: Atlantis [7] and Vadim Suraev implementation (no reference available, it could be consulted in the perforce server of MPLS-Linux project [145]).

The former aimed at implementing RSVP-TE as a signalling protocol for the MPLS architecture.

The second daemon developed by Vadim Suraev is more recent implementation of RSVP-TE and compatible with 2.6.x kernels. The interaction with Cisco routers and Quagga are being studied now.

Rsvp-Te Daemon: Atlantis

First of all the daemon supports the set up of Label Switched Paths (LSPs) in the network according to the IP routing tables or by explicitly specifying the hops to be traversed (shortest path LSPs and Explicitly Routed-LSPs respectively). There s an RSVP API (RAPI), an API to aid developers to build custom applications that interact with the RSVP daemon. There is support for DiffServ allowing differentiating the forwarding behaviour based on the value in the EXP field of the MPLS header. There is also support for IntServ where resources are explicitly allocated on a per-LSP level. Traffic can be mapped on the LSPs very flexibly based on the destination address, protocol, destination ports and port ranges of the IP packets.

There is also the ability to trace an LSP, comparable to IP s traceroute. It checks the route taken by an LSP by probing the routers along the path. Resilience is also supported with LSP rerouting and LSP protection switching.

The overall architecture consists of a number of components both in user space and kernel space. The important parts of the kernel that are used are netfiler to classify the packets, QoS and faire queuing to support differentiating between flows and of course MPLS support. The prime user space component is the RSVP daemon that is responsible for the RSVP signaling and the maintenance of the MPLS state. The daemon is responsible for the allocating and installation of the MPLS labels during LSP set up and freeing and removing labels on LSP tear down.

Two components use the RAPI: rtest and rapirecv. rtest is an application that takes LSP requests and issues them to the daemon. rapirecv is an application that receives label requests at the egress and dictates the daemon to send a response back to the ingress. rtest2 and rapirecv_auto (not shown in the figure) are extended version of rtest and rapirecv respectively that support the automatic set up of a (large) number of LSPs.

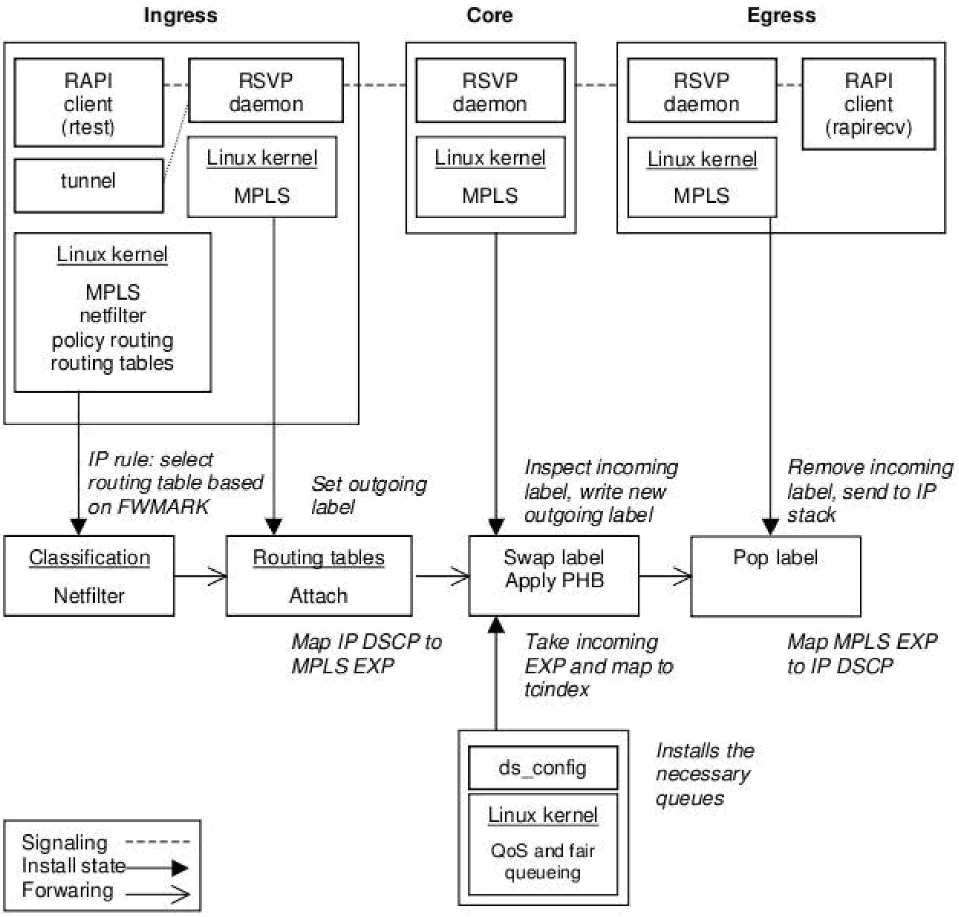


Figure 93: DiffServ over MPLS using RSVP-TE under Linux overall architecture

In the ingress the packets are classified with netfilter. Packets are filtered on the OUTPUT and PREROUTING chain of the mangle table. The mangle table is needed because the fwmark needs to be set. The OUTPUT and PREROUTING chains are used in order to filter on both locally generated and incoming traffic. Based on the value of the fwmark a routing table is selected (using policy routing). In the resulting routing table there is a MPLS tunnel interface acting as the default gateway. The tunnel interface encapsulates the packets on the LSP by attaching the correct outgoing label. The incoming DSCP is mapped to the EXP field in the MPLS header. A limitation of this architecture is that netfilter can only write (mark) on the mangle table and that only a single mark operation is possible. So while we do can map traffic on the LSP also setting the DSCP at the same node is impossible. The solution is to set the DSCP before the traffic enters the ingress LSP.

In the core node the MPLS stack inspects the incoming label and sets the new outgoing label and next hop. At the DiffServ level the current EXP value of the packet is inspected. There is a mapping from this EXP value to a tcindex. The tcindex in turn determines the correct outgoing queue so the correct forwarding behavior (PHB) can be applied.

Finally in the egress the incoming EXP field is mapped to the DSCP field and then the MPLS header is stripped off and the packet is sent to the IP layer.

Atlantis offers great opportunities for control plane emulation and has already been tested in several projects. For instance, it was possible with this tool to set-up LSPs between Linux boxes and commercial routers (Juniper and Cisco). Note however that the code is not very stable, particularly when the number of LSPs increases and that some limitations exists. Some of these limitations:

- No resource reservation. Even if the LSP required bandwidth field is transported, no resource reservations are provided. This is not really an issue if the user plane is not to be studied. However, this means that there is no interaction with OSPF-TE. For instance, in order to decrease the available link capacity once a LSP is set-up which is a problem even for control plane emulation.
- Atlantis has not supported MPLS hierarchy (label stack). This may become an issue for multilayer TE experimentations or in order to move towards a GMPLS emulated control plane.

MPLS-Linux Project

MPLS for Linux [145] is an open source effort to create a set of MPLS signalling protocols and an MPLS forwarding plane for the Linux operating system. So far a MPLS forwarding plane for the Linux 2.6.x kernel tree and an implementation of RFC3036 (LDP) has been created.

MPLS for Linux is made up of two projects:

- mpls-linux: MPLS forwarding for the Linux Kernel
- ldp-portable: A portable implementation of RFC3036

MPLS-Linux project from sourceforge offers a big opportunity to test MPLS in a Linux environment. However, it’s still under development, and many features, especially RSVP-TE, are not yet available.

Some reports have been found on Internet that have probed that a hybrid Cisco – Linux scenario with MPLS and OSPF-TE can be achieved. Quality of Service may be achieved in two ways:

- By QoS queues at ingress or egress interfaces
- By marking the EXP bits at the MPLS header

There is no need of choosing one of these possibilities. Both can be applied at the same time.

D.5. Commercial Implementations

Data Connection

Data Connection (http://www.dataconnection.com) Limited (DCL), along with its MetaSwitch division (www.metaswitch.com), is one of the world's leading telephony and communications technology companies.

Data Connection has led the development of MPLS control plane software for over 5 years, taking a leading role in driving the optical standards. Data Connection's MPLS software, DC-

MPLS, is a high quality, portable MPLS source code solution designed for the most demanding applications. DC-MPLS provides the MPLS label distribution protocols RSVP and LDP, is fully IPv6 compliant, and includes extensions for optical transport networks and VPNs.

The following block diagram shows the high level software architecture of the DC-MPLS protocol stack.

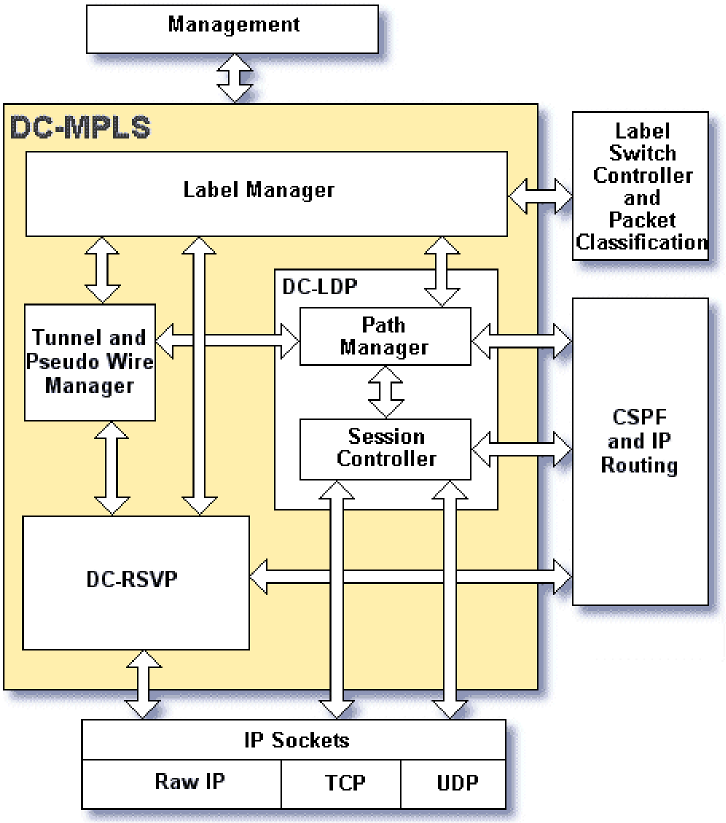


Figure 94: High level software architecture of the DC-MPLS protocol stack

DC-RSVP provides full support for the RSVP-TE protocol (RSVP Traffic Engineering). This includes

- Parsing and building of RSVP messages and objects
- Validation of received message semantics and syntax
- Entry points to support proprietary objects
- Maintenance of path, resv and session control blocks
- All refresh processing
- Interface to IP routing to obtain next-hop address (or two separate next-hop addresses for signalling and data links for out-of-band signalling), which may invoke a CSPF calculation
- Replication of state for high availability and hot software upgrade.

The DC-MPLS Label Manager is a common component that manages the LSR MIB and provides a single interface to the data plane for both RSVP-TE and LDP. This includes:

- Interfacing to resource management and switch programming for reserving and cross-connecting Label Switched Paths (LSPs)

- Optionally allocating labels (this can also be handled by OEM provided software)
- Full support for the LSR MIB, including providing an interface for management establishment of static LSPs
- Replication of state for high availability and hot software upgrade.

The Tunnel and Pseudo Wire Manager (TPM) component manages the MPLS Traffic Engineering (TE), Tunnel, and Pseudo Wire MIBs, including providing a management interface to establish LSPs at the ingress, and to respond to requests to establish LSPs at the egress. This includes

- Full read, write, create access to the MPLS TE, Tunnel, and Pseudo Wire MIBs
- Ability to add customized data to the MIBs
- Replication of state for high availability and hot software upgrade.

Integration Options

Data Connection's portable protocol software provides high performance, flexible and reliable signalling components for current and next-generation networking devices.

The software is extremely modular in design, with published interfaces: customizable macro interfaces to the operating system services and asynchronous, message-based interfaces between the components. This design enables the software to be readily integrated with any operating system and underlying hardware, as well as custom user applications, third-party components, and legacy systems.

Data Connection partners with third-party organizations to supply integrated solutions that combine each party's technology and minimize the integration effort and risk for its customers. In conjunction with its partners, Data Connection can provide an extensive range of example code and pre-integration solutions. This includes an out-of-the-box software router for Linux, which demonstrates integration with all the necessary complementary components to create a complete multi-protocol router.

Data Connection's portable protocol software products are available separately and as pre-integrated control plane solutions for different applications

Management Interface Integration

All the products use a MIB-like interface for management and configuration, which is designed for integration with CLI, GUI and SNMP management suites. They can also supply pre-integrated CLI, XML/Web and SNMP management interfaces using embedded MINDTM from its partner Silicon & Software Systems (S3).

Operating System Integration

All of the products share the same N-BASE portability environment, which ensures that they are fully portable. Pre-ported solutions are provided for popular embedded operating systems including VxWorks, Linux, Windows, OSE, Solaris and LynxOS. These standard ports may also be used as the model for other operating systems, almost every operating system and hardware platform available.

IP Stack Integration

The asynchronous, message-based interfaces to the underlying IP transport allow the products to be efficiently integrated in all system architectures. They provide integration with the standard IP stack for each operating system as part of their standard N-BASE ports, and with third-party IP stacks from our partners.

Data Plane Integration

For performance reasons, Data Plane functionality is normally closely integrated with the communications hardware, so most chipsets are supplied with their own Layer 1 & 2 drivers. Data Connection uses asynchronous message-based interfaces to the data plane and works closely with hardware manufacturers to ensure its interfaces map easily onto the underlying functionality.

IP Infussion

IP Infusion delivers an advanced software development platform that powers carrier and enterprise communication equipment. With a unique modular architecture and the industry's broadest suite of service protocols, IP Infusion enables product differentiation and market agility for many of the world's leading communication equipment vendors and service providers. Incorporated in Delaware in October 1999, IP Infusion is headquartered in San Jose, California, and is a wholly owned, independently operated subsidiary of ACCESS Co., Ltd.

MPLS approaching

IP Infusion's ZebOS Advanced Routing Suites (ARS) provides a series of control plane Software solutions for IPv4, IPv6, Multicast, MPLS-VPN, and Layer 2 protocols. Particularly for MPLS and Traffic Engineering (TE), IP Infusion provides TE extensions to OSPFv2 and IS-ISv4 as well as RSVP-TE, LDP, and CR-LDP signaling protocols plus MPLS forwarder software. In the area of DiffServ, IP Infusion has developed DiffServ extensions to RSVP-TE for supporting of E-LSP and L-LSP (RFC3270). Further more; IP Infusion has developed DiffServ extensions for support of MPLS Traffic Engineering “draft-ietf-tewg-diff-te-proto” in a DiffServ-aware MPLS-TE.

The ZebOS ARS Multiprotocol Label Switching (MPLS) modules provide a complete solution for the rapid integration of MPLS functionality into enterprise, edge, and core communications equipment. The ZebOS ARS supports both MPLS Traffic Engineering (TE) and MPLS Virtual Private Network (VPN) capabilities, as well as extensions for the support of DiffServ (Differentiated Services) and DiffServ with Traffic Engineering (DiffServ-TE).

The ZebOS MPLS modules support the following protocols and features:

- Resource Reservation Protocol-Traffic Engineering (RSVP-TE) with Graceful Restart
- Label Distribution Protocol (LDP) with Graceful Restart and Control Word Signaling
- Constrained Shortest Path First (CSPF for OSPF and IS-IS)
- MPLS-VPN using Border Gateway Protocol (BGP) VPN extensions (RFC 2547)
- Virtual Private LAN Service (VPLS with Martini)
- DiffServ and DiffServ with TE Extensions
- Pseudowire (PW) support for Layer 2 VPN

D.6. Traffic Engineering

Traffic Engineering (TE) extensions use the RSVP and LDP dynamic-signaling protocols to communicate to the ZebOS MPLS Forwarder or a third-party MPLS forwarder. The RSVP-TE extension enables MPLS to scale into large and complex IP-based communications equipment.

The ZebOS ARS also supports the Constrained Shortest Path First (CSPF) algorithm. Using a TE database and pre-existing Label Switched Paths (LSPs), the CSPF module calculates — on demand — an optimum Explicit Route (ER), based on the specified constraints. The resulting ER is used by a signaling protocol, either RSVP-TE or LDP, to set up LSPs.

ZebOS ARS MPLS Suite

MPLS Forwarder – Because ZebOS ARS supports IPv4, IPv6 and MPLS networks, it is well positioned not only to support core, edge, and access routing and switching platforms, but also to act as the standard routing protocol for an entire range of IPv6-enabled devices, including Small Office Home Office (SOHO) gateways; wireless, access, and security devices; and devices that support VPNs and Voice-over-Internet Protocol (VoIP) technology, which require Quality of Service (QoS) and superior bandwidth management. All ZebOS protocols use the Network Services Manager (NSM) to update Information databases and to interface with the Forwarder.

RSVP-TE – The Resource ReSerVation Protocol (RSVP) is a signaling protocol that supports explicit routing capability. To do this, a simple Explicit Route (ER) object is incorporated into RSVP PATH messages. The object encapsulates a sequence of hops, which constitute the explicitly routed path. Using this object, the paths taken by the label-switched RSVP-MPLS flows can be pre-determined without conventional IP routing. The explicitly routed path can be administratively specified or computed based on CSPF and policy requirements dictated by the operator through the trunk node. A valuable application of explicit routing is Traffic Engineering (TE). Using explicitly routed LSPs, an ingress node can control the path through which traffic flows from itself, through the MPLS network, to the egress node. Explicit routing is therefore useful for the optimization of network resources and an increase in the quality of traffic oriented performance.

LDP – The Label Distribution Protocol (LDP) is a protocol in which two label-switched routers (LSR) exchange label mapping information. The two LSRs are called LDP peers and the exchange of information is bi-directional. LDP is used to build and maintain databases of LSR that are used to forward traffic through MPLS networks. LDP generates labels for, and exchanges labels between, peer routers. It works with other routing protocols (RIP, OSPF and BGP) to create the label-switched paths (LSP) used when forwarding packets. An LSP is the path taken by all packets that belong to the Forwarding Equivalence Class (FEC) corresponding to that LSP. This is analogous to establishing a virtual circuit in ATM (Asynchronous Transfer Mechanism). In this way, ZebOS LDP assigns labels to every destination address and destination prefix provided by ZebOS. The LDP interface to the MPLS forwarder adds labels to, and deletes labels from, the forwarding tables.

DiffServ – Differentiated Services (DiffServ) is a networking architecture that specifies a simple, scalable and coarse-grained mechanism for classifying and managing network traffic, and providing Quality of Service (QoS) guarantees for service providers. DiffServ employs a sophisticated policy to determine how to forward network data, so it is more advanced than earlier Quality of Service (QoS) or Type of Service (ToS) protocols. In order to provide a flexible DiffServe-over-MPLS solution, the ZebOS DiffServ module is available as an extension to the ZebOS RSVP-TE module. The ZebOS DiffServ module enables network traffic to be specified and prioritized by class, so that certain kinds of traffic, for example voice traffic, attain precedence over other types of traffic.

DiffServ-TE – Since DiffServ by itself lacks the ability to efficiently use network transmission resources, IP Infusion also offers a DiffServ-aware MPLS Traffic Engineering (DiffServ-TE) module compliant with RFC 4124. The ZebOS ARS DiffServ-TE module performs traffic engineering per DiffServ class, rather than at an aggregate level. By combining DiffServ with MPLS-TE, routing devices can simultaneously classify and prioritize traffic, and achieve fined-grained optimization of transmission resources.

GateD (Nexthop)

The GateD (http://www.nexthop.com) suite of routing and MPLS software is a complete control-plane solution, with all requisite protocols packaged conveniently for inclusion in a variety of networking devices. Available GateD protocols include RIP, OSPF, BGP, IS-IS, DVMRP, IGMP, PIM-SM, PIM-SSM, PIM-DM, MSDP, MP-BGP for IPv6, IS-IS for IPv6, OSPFv3, MPLS, VRE (Virtual Routing Environment) and VPN (layer 3 MPLS-BGP virtual private networking) support. Equipment manufacturers can pick and choose from these individual components or license a variety of packages targeted at specific applications.

NextHop's MPLS solution is the first complete MPLS package tightly integrated with IP routing for enhanced scalability and performance. The MPLS solution includes a comprehensive stack with label manager and interface manager, as well as both LDP and RSVP-TE signaling. GateD AMI provides a consistent interface for management of both routing and MPLS protocols, and a complete CLI is available. In addition to the scalability enhancement inherent in an integrated routing and MPLS solution, GateD MPLS can be used in conjunction with NextHop's IGPs for a fast, complete traffic engineering offering, and it is pre-integrated with BGP for support of industry standard layer-3 VPNs (with the GateD NGC Virtual Routing Environment). The integration of all parts of the control plane allows for quicker porting and integration, as well as enhanced transaction speed and general scalability.

MPLS approaching

The foundations of NextHop MPLS lay in the shared components of the Label Manager, the TE Tunnel Manager, and the Interface Manager. Together, these components provide a clear interface to the data plane, manage LSPs, and interact with the Layer 3 IP routing protocols. The Label Manager allocates and manages labels for use in a variety of MPLS applications, including most significantly traffic engineering, and layer 2 and layer 3 virtual private networking. As appropriate, the Label Manager interacts directly with the IP routing protocols (for example, when using BGP as a label distribution protocol). The TE Tunnel Manager is the nexus of the label distribution protocols, traffic engineering information (such as available bandwidth or link color), and the constraint based routing algorithms of the interior gateway protocols (IS-IS and OSPF).

Finally, the Interface Manager is the link back from MPLS to the routing protocols. It allows for functionality such as IGP shortcut and forward adjacency, where OSPF or IS-IS can run directly over a traffic-engineered tunnel, allowing network operators to optimize the flow of routing data over their core optical networks either in conjunction with or in migration away from classical ATM full-mesh topologies.

LDP and RSVP-TE are the “on-the-wire” workhorses of any MPLS implementation. These are the signaling protocols that allow for the creation of label-switched paths across the network. More than basic implementations of standard RFCs, these label distribution protocols are fully integrated with NextHop's complete MPLS and routing solution, so that they interoperate with static LSPs, can create tunnels over tunnels (e.g. LDP over RSVP), and can implement advanced features, such as fast reroute, seamlessly.

As with all NextHop routing products, the NGC MPLS strikes the perfect balance between richness of features, stability, and scalability. The NGC product line is designed to meet the feature and scalability requirements of core, edge and aggregation boxes target for carrirers' networks over the next five years. And like NextHop classic routing offerings, MPLS is fully tested for conformance and interoperability.

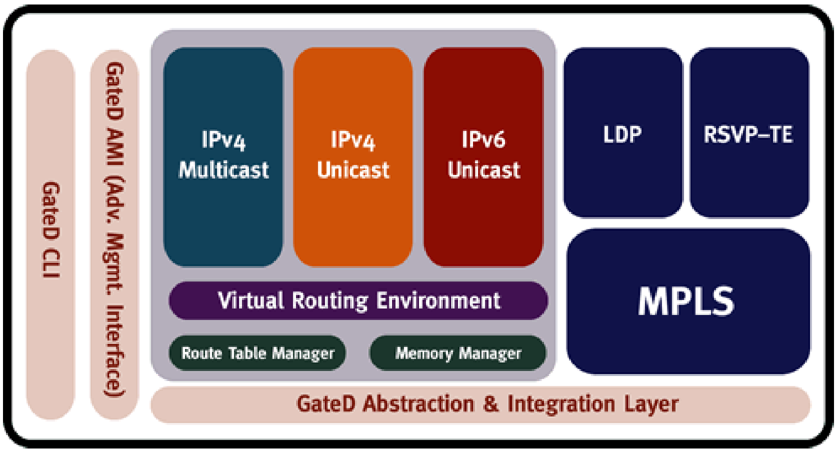

GateD CLI
GateD AMI (Adv. Mgmt. Interface)
IPv4 Multicast
IPv4 Unicast
IPv6 Unicast
LDP
RSVP-TE
Virtual Routing Environment
Route Table Manager
Memory Manager
MPLS
GateD Abstraction & Integration Layer

Figure 95: GATED architecture

Supported Protocol RFCs and Drafts

- Label Distribution Protocol (RFC 3036)
- RSVP (RFC2205)
- RSVP-TE (RFC 3209)

Supported Other RFCs

- Graceful Restart Mechanism for Label Distribution Protocol (RFC 3478)
- Generalized Multi-Protocol Label Switching (GMPLS) signaling Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions (RFC 3473 - section 9)
- Extensions to GMPLS RSVP Graceful Restart
- Multi protocol Label Switching Architecture (RFC3031)
- MPLS Label Stack Encoding (RFC 3032)
- Pseudo Wire Setup and Maintenance using LDP
- MPLS support of Differentiated Services (RFC 3270)
- Fast Reroute Extensions to RSVP-TE for LSP tunnels (RFC 4090)

Aricent

Aricent MPLS (Multi Protocol Label Switching) software is an efficient and portable implementation of the label switching capability mechanism as required by industry standards. It supports RSVP extensions for traffic engineering (RSVP-TE), Label Distribution Protocol (LDP) with Constraint-based Routed Label Switched Paths (CR-LSP), MPLS Forwarding Module and Diff-Serv aware signalling. Aricent MPLS can be configured to enable Layer 2 Virtual Private Networks (L2-VPN) services.

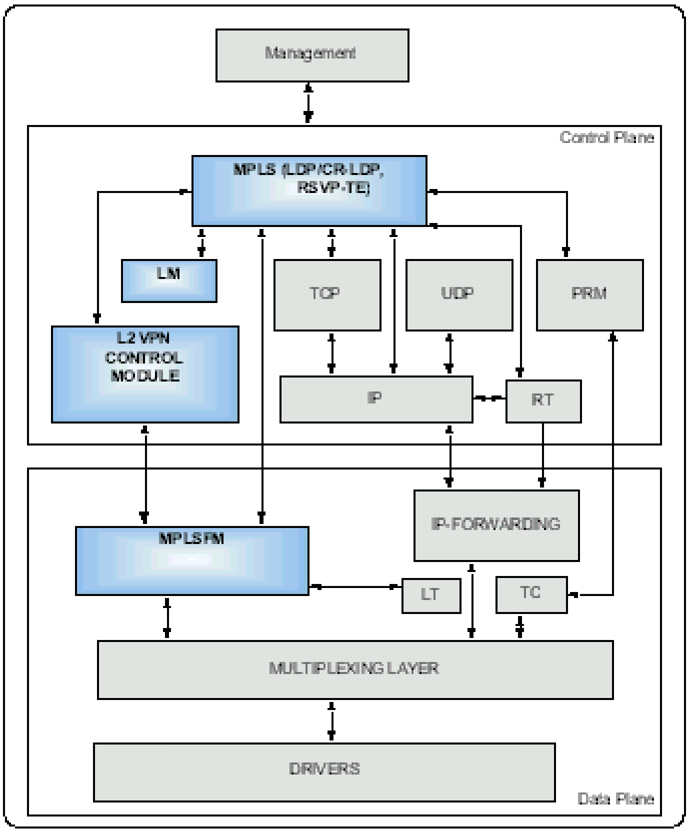


Figure 96: Aricent block diagram

Features supported

- Configurable as LER/PE or LSR.
- Manageable through SNMP and CLI.
- Software forwarding module included.
- Generic API support for using hardware based MPLS Forwarding Module implementation.
- Conforms to Future Software Architecture for Portability Release (FSAP2), thus ensuring highly portable code, which uses flexible buffer and timer management libraries
- Pre-integrated with Aricent IP and Aricent SNMP.

Generic Signaling Features

- Conservative Label Retention mode.

- RSVP-TE and CR-LSP tunnel modifications.
- Ethernet and ATM interface.
- Multiple FEC classifications based on IP Address Prefix
- Host Address

RSVP-TE features

- Traffic Engineering (TE) tunnels.
- Downstream on Demand label distribution mode.
- LSP Pre-emption (Set up and Holding).
- Resource Class Affinity Attribute handling.
- Provides Loop detection using RRO.
- Hello message extensions for rapid detection of node failures.
- RSVP Refresh Overhead Reduction and Reliable Message Delivery.
- Diff-Serv aware E-LSP and L-LSP signalling.

LDP/CR-LDP features

- Basic and extended Discovery mechanisms.
- Downstream On Demand label distribution mode.
- Ordered and Independent Control.
- Loop detection using path vectors and hop counts.
- Penultimate Hop Popping (PHP).
- Stacked LSPs.

LDP features

- Downstream Unsolicited label distribution mode.
- Liberal Label Retention Mode.
- VC/VP Merge Capability.

CR-LDP features

- Constraint-based routed LSP (CR-LSP).
- Diff-Serv aware E-LSP and L-LSP signaling.
- LSP Preemption (Setup and Holding).

L2-VPN features

- Control Plane support for Virtual Private Wired Services (VPWS)).
- Pseudowire Management and Maintenance.
- Control plane layer 2 service support for Ethernet including VLAN
- Control plane support for collecting and maintaining PW-VC Informational statistics.

Packaging options

- Aricent MPLS - LM, TE-MIB, LDP, CR-LDP, RSVP-TE and FM
- Aricent MPLS with LDP - LM, TE-MIB, LDP, CR-LDP and FM
- Aricent MPLS with RSVP-TE - LM, TE-MIB, RSVP-TE and FM
- Aricent MPLS with L2-VPN - LM, TE-MIB, LDP, CR-LDP, RSVP-TE and FM

D.7. Conclusions

Routing Protocol Controller

Quagga with extensions of OSPF-TE is the only open routing software with some compatibility with MPLS. There are other big projects of open source routing like XORP [218] but now it has not and neither intends to have compatibility with MPLS.

Forwarding Engine (user plane) MPLS/Linux project developed by James R. Leu. The major problem of MPLS/Linux is the lack of documentation, so the entry barrier to kernel hackers that want to know the inner and implementation details is too high. This issue and the inherent problem of have very few developers to the project makes that the evolution of it be slow and hard.

Signalling Protocol Controller

Atlantis, it is actually not stable at all, for instance when the number of LSPs increases. Furthermore there are no resource reservations. Even if the LSP required bandwidth field is transported, no resource reservations are provided. This is not really an issue if the user plane is not to be studied. However, this means that there is no interaction with OSPF-TE for instance in order to decrease the available link capacity once a LSP is set-up which is a problem even for control plane emulation.

Atlantis has not supported MPLS hierarchy (label stack). This may become an issue for multi-layer TE experimentations or in order to move towards a GMPLS emulated control plane.

The other signalling protocol that is available is Vadim's implementation, but it's currently studied and some integration problems have been issued.

The following blocks still missing:

- MPLS-TE Connection Controller
- MPLS-TE Routing Controller
- Management and Policy Client
- The interaction between these blocks is also to be worked on (it's the major problem). In particular the interaction between RSVP-TE and OSPF-TE in case of bandwidth reservations.

On the other hand in commercial implementations, as we have seen before, there are a lot of the last RFCs and drafts that IEEE and other organizations like IETF developed. The main cause why Open Source implementations don’t evolve quickly is that there aren’t enough developers and efforts; for instance, MPLS-Linux: that is the most important project to bring a Linux Router the capacity of act as MPLS router, has only one developer and the future for open source implementations of MPLS is not very clear.

The problem that educational researchers had is that the prize of implementation a commercial resource is prohibitive. For example our team contact with vendors described before and the response that we got was that instead we are educational institute, they need a minimum fee of around 50.000 Euros to catch their interest. All their processes are setup for substantial customer engagements.

Appendix E: MPLS-TE optimization with HGA

In this appendix preliminary results of an incipient research on optimizing MPLS-TE networks with hybrid genetic algorithms are provided. This testing has been done by means of simulations.

This first approach for testing the HGA interaction with MPLS-TE has been done over the TOTEM platform for traffic engineering simulations. This platform has been developed by the TOTEM project [196] of the Université Catholique de Louvain and the Université de Liège, both in Belgium.

The objective of the TOTEM project (TOolbox for Traffic Engineering Methods) is to develop a toolbox of algorithms for traffic engineering purposes. Therefore, they unify the algorithms which have already proposed these last years and develop new techniques of traffic engineering. These techniques take into account the distribution of the traffic, the fault-tolerance requirement and the support of the quality of service. They develop generic algorithms for the optimization of networks of big size which apply, on one hand, to IP networks, and on the other hand, to networks operated with (G)MPLS. Some of these algorithms require extensions to the routing (OSPF-TE, ISIS-TE, BGP) or signalling (RSVP-TE) protocols. The toolbox is available in open source and designed such that its elements can easily be integrated in various platforms such as Linux PCs, routers, and open source network simulators like NS and/or J-Sim.

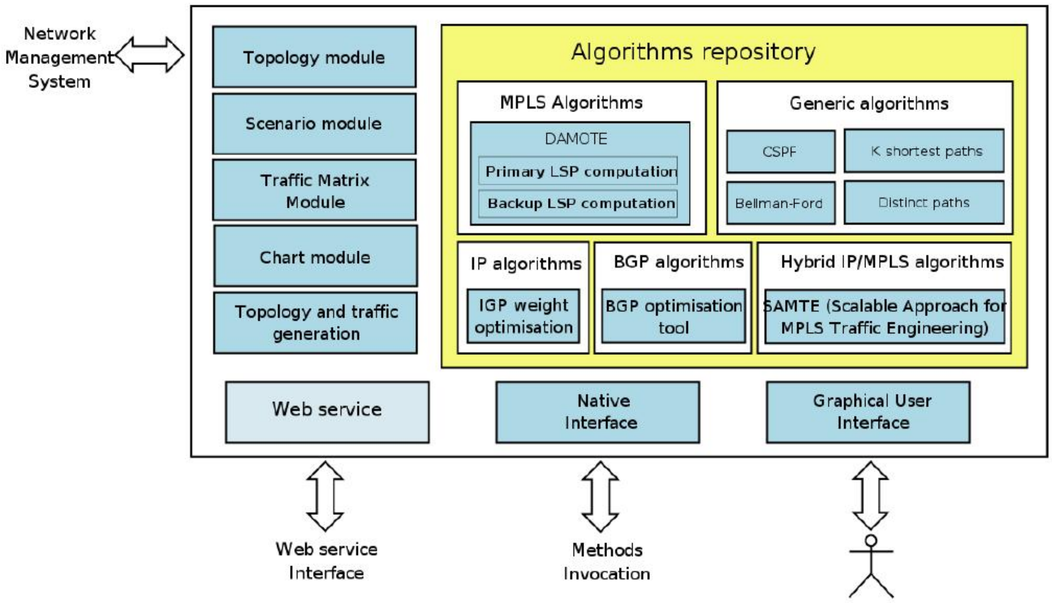


Figure 97: TOTEM architecture

There have been used 5 different scenarios with different optimization techniques.

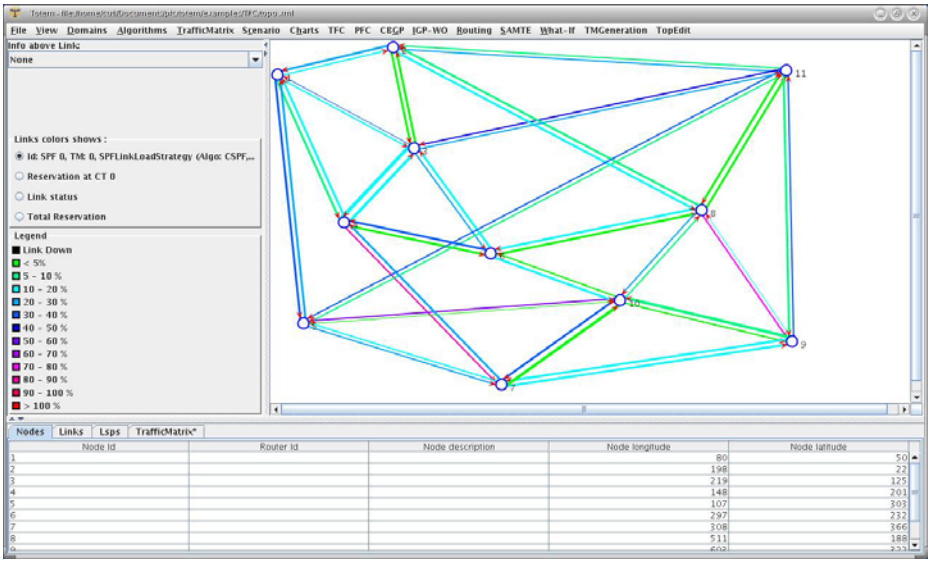
- **Scenario 1**: IGP simple rountig based on OSPF (with ECMP). Provided by TOTEM
- **Scenario 2**: OSPF routing appling an HGA algorithm, the same algorithm presented in Chapter 6 (with ECMP)
- **Scenario 3**: MPLS routing based in the minimum hop technique for the LSP generation. Provided by TOTEM
- **Scenario 4**: MPLS routing applying an HGA algorithm, the same algorithm presented in Chapter 6 (with ECMP).
- **Scenario 5**: SAMTE (hybrid IP/MPLS routing) applying an HGA algorithm, the same algorithm presented in Chapter 6 (with ECMP).

Two diffent networks have been used to do the simulations. The well known N11 used in chapter 6 for the HGA routing optimization and the Abilene network.

Simulations on the N11 network

The N11 network has been show in Figure 80 and has 11 nodes and 40 unidirectional links. The preliminary tests have compared diffent proposals for routing optimization over the same N11 network.

The traffic matrix and the link capacity matrix used have been provided by the TOTEM project in order to accelerate the firts results and are different from those used in Chapter 6. Therefore direct comparison cannot be established yet.



Info above Link: None

Links colors shows :

● Id: SPF 0, TM: 0, SPFLinkLoadStrategy (Algo: CSPF,...

○ Reservation at CT 0

○ Link status

○ Total Reservation

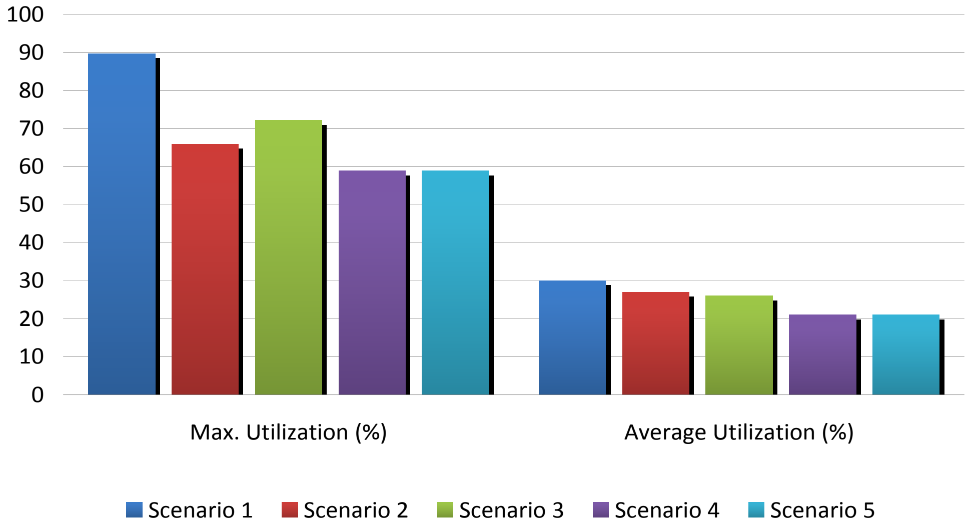
Legend

■ Link Down
■ < 5%
■ 5 - 10 %
■ 10 - 20 %
■ 20 - 30 %
■ 30 - 40 %
■ 40 - 50 %
■ 50 - 60 %
■ 60 - 70 %
■ 70 - 80 %
■ 80 - 90 %
■ 90 - 100 %
■ > 100 %

Node Id	Router Id	Node description	Node longitude	Node latitude
1			80	50
2			198	22
3			219	125
4			148	201
5			107	303
6			297	232
7			308	366
8			511	188
9			602	222

Figure 98: Example of TOTEM application

The first results are presented in the next graph.



The best results are obtained in scenario where HGA is applied, and with special attention to the 4th and 5th MPLS scenarios.

The regular use of the network may be enhanced by the use of these techniques passing from the 89% usage of most congested link using OSPF to 66% using the HGA algorithm, and from 72% using MPLS to 59% using MPLS-TE.

The average utilization value provides the average use of the links in the overall network and it can be seen a great improvement with MPLS-TE with genetic algorithms, passing from an average of 30% to 21%.

These preliminary results prove that it is a valid line of research.

Furthermore, it has been made an analysis in order to determine if a higher greater number of individuals in the genetic algorithm would lead to a better solution. The next graph shows how the improvement in the solution is not significant in contrast with the high increase in the execution time.

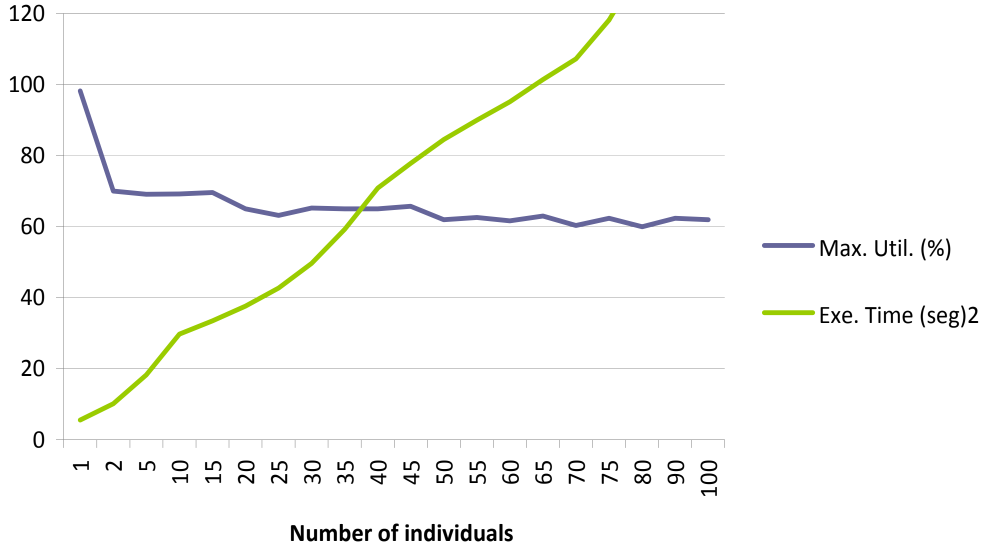


Figure 99: Relation between Max. Utilization vs. Execution Time

Simulations on the Abilene network

The Abilene network and has 11 nodes and 28 unidirectional links and is already implemented in the TOTEM tool. The preliminary tests have compared diffent proposals for routing optimization over the same Abilene network. The traffic matrix and the link capacities have been provided by the TOTEM tool. In this case the traffic matrix is much complex than the previous one and there are many sources going to many destinations.

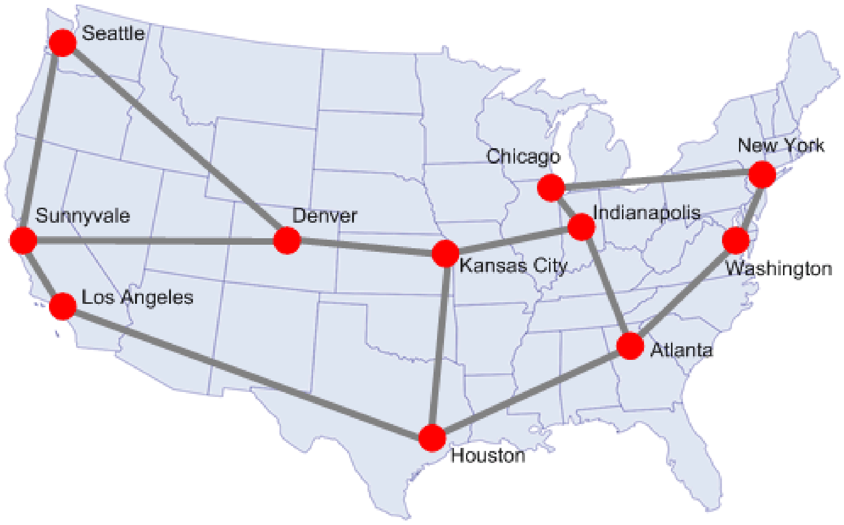
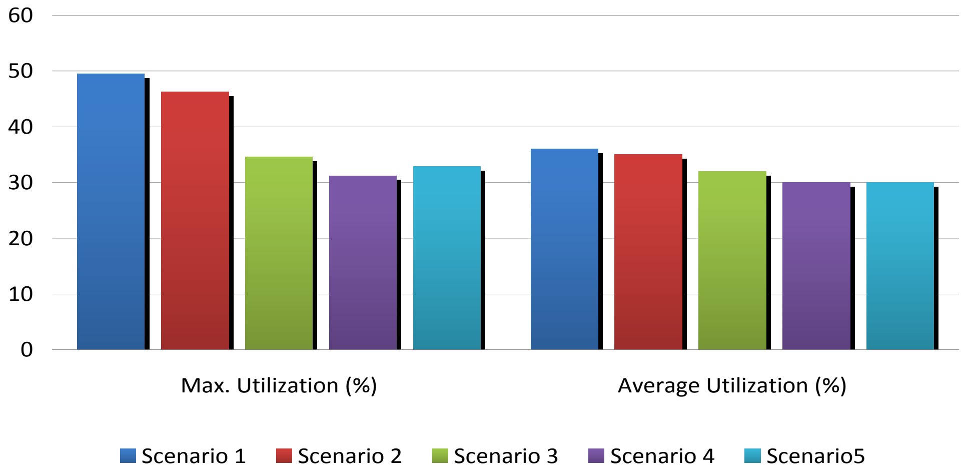


Figure 100: Abilene network

The first results are presented in the next graph.



The best results are obtained with MPLS and more especially in those scenarios where HGA is applied, and with special attention to the 4th and 5th MPLS scenarios.

The maximum use of a link drops to 31% from an initial of 49%.

The average use of the network lowers from 36% to 30% with MPLS-TE using HGA.

These preliminary results confirm that it is a valid line of research.

Universitat Ramon Llull

Aquesta Tesi Doctoral ha estat defensada el dia ____ d __________________ de ____

al Centre ____

de la Universitat Ramon Llull

davant el Tribunal format pels Doctors sotasignants, havent obtingut la qualificació:

President/a

____

Vocal

____

Vocal

____

Vocal

____

Secretari/ària

____

Doctorand/a

____