Torsten Braun, Michel Diaz, José Enríquez-Gabeiras, and Thomas Staub

End-to-End Quality of Service Over Heterogeneous Networks

February 12, 2008

Springer

Foreword

Preface

The Internet has evolved from an academic network for data applications such as file transfer and net news, to a global general-purpose network used for a variety of different applications covering electronic mail, voice over IP, television, peer-to-peer file sharing, video streaming and many more. The heterogeneity of applications results in rather different application requirements in terms of bandwidth, delay, loss, etc. Ideally, the underlying network supports such Quality-of-Service parameters such that applications can request the desired services from the network, and do not need to take actions by themselves to achieve the desired communication quality. Initially, the Internet was not designed to support Quality-of-Service, and only since the last decade have appropriate mechanisms been developed. Those mechanisms mainly operate on the Internet Protocol (IP) level, but also network-specific mechanisms—e.g., targeted to particular wired/wireless access network technologies—are required.

The goal of the European 6th Framework Programme (FP6) Integrated Project ”End-to-end Quality of Service Support over Heterogeneous Networks” (EuQoS) was to develop, implement, and evaluate concepts and mechanisms to support QoS end-to-end, meaning that QoS mechanisms in end systems, access networks, inter-domain links and within domains must be supported. The EuQoS project developed an impressive set of innovative solutions and novel scientific ideas to support end-to-end QoS in the Internet. New mechanisms and concepts were designed and implemented in a European-wide distributed testbed. In addition to the rather technical design and implementation work, the project also developed training material introducing basic QoS mechanisms and techniques. Several e-learning modules were developed and are currently being used at several partner universities for teaching on MSc or PhD levels.

The significant technical and educational results achieved during the EuQoS project, motivated us to use the gained knowledge and experiences of the project partners and write this book on end-to-end QoS in heterogeneous IP networks. The book basically consists of three parts. In Chapters 1-4, we discuss QoS mechanisms and protocols such as scheduling schemes, QoS architectures metrics and measurement techniques, traffic engineering and signalling protocols, and the latest standardisation activities. Chapter 5 describes related work and recent development in the area of transport protocols, in particular how TCP can be optimised towards QoS support and fairness. The EuQoS system presented in Chapter 6 extends and combines the basic mechanisms discussed in the previous chapters. We show how a combination of different QoS enabling mechanisms and protocols can be used and extended to build a comprehensive end-to-end QoS architecture over heterogeneous wired/wireless access networks. To evaluate QoS mechanisms and architectures, appropriate evaluation schemes are required. The two chapters in the annex describe how simulation—in particular the well-known network simulator ns-2—as well as emulation techniques can be used for tests and evaluations.

This book, which is based on the achievements of the EuQoS project, would not have been possible to compile without the funding from the European Commission, as well as the tremendous efforts and enthusiasm of all the people involved in the project. Special thanks to Mark Günter for proof-reading the text contributions to this book.

Torsten Braun
Michel Diaz
José Enriquez Gabeiras
Thomas Staub

Bern, Toulouse, Madrid. January 2008.

Acknowledgements

The book editors and authors would like to thank all people who were involved in the EuQoS project:

- **Telefonica I+D:** José Enríquez Gabeiras, Francisco Javier Ramón Salguero, Gerardo García de Blas, Antonio J. Elizondo Armengol, Francisco, Romero Bueno, Jesús Bravo lvarez, Jorge Andrés Colás, María Ángeles, Callejo Rodríguez, María Luisa García Osma
- **University of Pisa (CPR/UoPisa):** Enzo Mingozzi, Giovanni Stea, Luciano Lenzini, Luca Bisti, Claudio Cicconetti, Linda Martorini, Abraham Gebrehiwot, Simone Bisogni, Paolo Sozzi
- **Elsag Datamat:** Enrico Angori, Giuseppe Martufi, Marco Carusio, Alessandro Giorgini, Andrea Paselli, Giovanni Saccomandi, Marco Mauro
- **CNRS (LAAS-CNRS, ENSICA):** Michel Diaz, Florin Racaru, Ernesto Exposito, Philippe Owezarski, Patrick Senac , Christophe Chassot , Nicolas Larrieu, Laurent Dairaine, Mathieu Gineste, Nicolas Van Wambeke, Slim Abdellatif, Sébastien Ardon, Roberto Willrich, Guillaume Auriol, Silvia Farraposo
- **France Telecom R&D:** Olivier Dugeon, Walid Htira, Michel Bourbao, Pascal Le Guern, Jean-Louis Le Roux, Stéphane Statiotis, Régis Fréchin, Claire Teisseire
- **Polska Telefonia Cyfrowa (ERA):** Michal Obuchowicz, Robert Parzydo , Adam Flizikowski, Edyta Rafalska, Karol Jez, Krzysztof Horszczaruk, Krzysztof Bronarski, Krzysztof Samp, Maciej Rozowicz, Michal Dudzinski, Pawel Caban, Piotr Zadroga, Slawomir Tkacz
- **Martel:** Martin Potts, Mark Guenter, Sandra Wittwer
- **NICTA:** Emmanuel Lochin, Guillaume Jourjon, Sebastien Ardon, Ernesto Exposito, Feiselia Tan, Laurent Dairaine
- **PointerCom:** Roberto Marega, Stefano Salsano, Donald Papalilo, Gianluca Martiniello, Valeria Calcagni
- **Polish Telecom R&D:** Zbigniew Kopertowski, Jaroslaw Kowalczyk, Tomasz Ciszkowski

- **Portugal Telecom Inovação (PTIN):** Jorge Carapinha, Nuno Carapeto, Paulo Loureiro, Arnaldo Santos, Eduardo Silva, Fernando Santiago, Helena Paula Matos, Hugo Manaia, Isabel Borges, Jacinto Barbeira, Filipe Peixinho.
- **Red Zinc:** Donal Morris, Brian Widger, Diarmuid O Neill, Léa Compin, Oscar Cuidad
- **Silogic:** Laurent Baresse, Benoît Baurens, Jean-Philippe Darmet, Yannick Lizzi, François Meaude
- **INDRA (previously SOLUZIONA):** Ignacio Fresno, Jaime Orozco, Luis Collantes Abril, Pablo Vaquero Barbón, Rubén Romero San Martín, Jorge Alonso, Maria Lurdes Sousa, Raul Manzano Barroso
- **Telscom AG:** Sathya Rao, Marcin Michalak
- **Technical University of Catalonia (UPC):** Jordi Domingo-Pascual, Loránd Jakab, Marcelo Yannuzzi, René Serral-Gracià, Xavier Masip-Bruin
- **University of Bern:** Torsten Braun, Thomas Staub, Dragan Milić, Marc Brogle, Marc-Alain Steinemann, Thomas Bernoulli, Gerald Wagenknecht, Markus Wulff, Patrick Lauer, Markus Anwander, Matthias Scheidegger
- **University of Paderborn/C-LAB:** Isabell Jahnich, Achim Rettberg, Chris Loeser, Michael Ditze, Kay Klobedanz, Sebastian Seitz, Andreas König, Volker Spaarmann, Matthias Grawinkel
- **University of Rome:** Antonio Pietrabissa, Francesco Delli Priscoli, Sabrina Giampaoletti, Emiliano Guainella, Erasmo Di Santo, Gianfranco Santoro, Ilaria Marchetti, Massimiliano Rossi
- **Universidade de Coimbra:** Edmundo Monteiro, Luís Cordeiro, Bruno Carvalho, Fernando Boavida, Gabriela Batista Leão, Isidro Caramelo, Jian Zhang, Jorge Sá Silva, Marilia Curado, Maxwel Carmo, Paulo Simões, Romulo Ribeiro, Vitor Bernardo, David Palma, Rui Vilão, Luís Conceição
- **Warsaw University of Technology:** Wojciech Burakowski, Andrzej Beben, Halina Tarasiuk, Jaroslaw Sliwinski, Jordi Mongay Batalla, Marek Dabrowski, Piotr Krawiec, Robert Janowski
- **Ericsson:** Antoine de Poorter, Julio López Roldan, Miguel Angel Recio, Jesus Renero Quintero, José Luis Agundez
- **Hospital Divino Espirito Santo:** José Manuel Ponte, António Vasco Viveiros, Carlos P. Duarte, Paula Maciel, José M. Jesus Silva, Maura Medeiros, Maria Dulce Raposo

Contents

Acronyms

The following list contains acronyms used in the book and their explanation. Most acronyms can be found in the index as well together with a page reference.

ALM	Application layer multicast
API	Application programming interface
CIDR	Classless Internet domain routing
COPS	Common Open Policy Service
DVMRP	Distance-vector multicast routing protocol
IGMP	Internet group management protocol
IP	Internet protocol
IPTV	Internet Protocol Television
IPv4	Internet protocol version 4
IPv6	Internet protocol version 6
ISP	Internet service provider
MM	Multicast Middleware
MOSPF	Multicast open shortest path first
NSIS	Next Step In Signalling
P2P	Peer-to-peer
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PIM	Protocol-independant multicast
QoS	Quality of Service
SDP	Session Description Protocol
SE	Signalling Entities
SIP	Session initiation Protocol
SSQ	Synchronize State Query
TCP	Transmission control protocol
TTL	Time-To-Live
UAC	User Agent Client
UAS	User Agent Server
UDP	User datagram protocol

VLSM	Variable length subnet mask
MPLS	Multi Protocol Label Switching
TE	Traffic Engineering
RIP	Routing Information Protocol
IGP	Interior Gateway Protocol
OSPF	Open Shortest Path First
IS-IS	Intermediate System-Intermediate System
RSVP	ReSerVation Protocol
IETF	Internet Engineering Task Force
FEC	Forwarding Equivalence Class
LSR	Label-Switching Router
LFIB	Label Forwarding Information Base
ATM	Asynchronous Transfer Mode
DLCI	Data-Link Connection Identifier
VPI	Virtual Path Identifier
VCI	Virtual Channel Identifier
BGP	Border Gateway Protocol
LDP	Label Distribution Protocol
LIB	Label Information Base
LSP	Label-Switched Path
CBR	Constraint-Based Routing
TED	Traffic Engineering Database
CSPF	Constrained Shortest Path First
SPF	Shortest Path First
CR-LDP	Constraint-based Routing Label Distribution Protocol
TSPEC	Traffic Specification
ERO	Explicit Route Object
BA	Behavior Aggregate
PHB	Per Hop Behavior
DSCP	Diff-Serv Codepoint
OA	Ordered Aggregate
PSC	PHB Scheduling Class
AF	Assured Forwarding
E-LSP	EXP-Inferred-PSC LSP
L-LSP	Label-Only-Inferred-PSC LSP
BE	Best Effort
DS-TE	Diff-Serv-aware Traffic Engineering
CT	Class Type
BC	Bandwidth Constraint
MAM	Maximum Allocation Bandwidth Constraints Model
RDM	Russian Doll Bandwidth Constraints Model
AC	Access Category
ADSL	Asymmetric DSL
AIFS	Arbitrary Inter-Frame Space
AP	Access Point

AS	Autonomous Systems
ASPB	AS-path Builder
ATM	Asynchronous Transfer Mode
BR	Border Router
BRAS	Broadband Remote Access Server
BRPC	Backward Recursive Path Computation
CAC	Connection Admission Control
CBR	Constant Bit Rate
CoS	Class of Service
CPE	Customer Premises Equipment
CRA	Continuous Rate Assignment
CW	Contention Window
DAMA	Demand Assignment Multiple Access
DCF	Distributed Coordination Function
DSL	Digital Subscriber Line
DSLAM	Digital Subscriber Line Access Multiplexer
DVB-S	Digital Video Broadcasting - Satellite
DVB-RCS	Digital Video Broadcasting - Reverse Channel Satellite
e2e CoS	End-to-end Class of Service
EDCA	Enhanced Distributed Coordination Access
ER	Edge Router
ES	Ethernet Switch
FCA	Free Capacity Assignment
FTP	File Transfer Protocol
GGSN	GPRS Gateway Support Node
HTD	High Throughput Data
IPLR	IP Packet Loss Ratio
IPTD	IP Packet Transfer Delay
IPDV	IP Packet Delay Variation
MAC	Medium Access Control
MT	Mobile Terminal
NCC	Network Control Centre
NRT	Non Real Time
OGGSN	Open GPRS Gateway Support Node
PCC	Path Computation Client
PCE	Path Computation Element
PCEP	PCE Protocol
PQ	Priority Queuing
PR	Peak Rate
RA	Resource Allocator
RBDC	Rate Based Dynamic Capacity
RM	Resource Manager
RNC	Radio Network Controller
RT	Real Time
SHDSL	Symmetrical High Bitrate DSL

SLA	Service Level Agreement
ST	Satellite Terminal
STD	Standard
TERO	Traffic Engineering and Resource Optimization
TOS	(Type of Service)
UTRAN	UMTS Terrestrial Radio Access Network
VBDC	Volume Based Dynamic Capacity
VBR	Variable Bit Rate
VDSL	Very High Bitrate DSL
VoD	Video on Demand
VoIP	Voice over IP
VTC	Video Teleconference
WFQ	Weighted Fair Queueing
WMM	WiFi Multi-Media
WRED	Weighted Random Early Detection
WRR	Weighted Round-Robin
CLI	Command Line Interface
EQ-BGP	Enhanced QoS Border Gateway Protocol
QoS NLRI	QoS Network Layer Reachability Information
DoP	Degree of Preference
SNMP	Simple Network Management Protocol
TMN	Telecommunications Management Network
SAAA	Security, Authentication, Authorization and Accounting
QoSR	Quality of Service Routing
xDSL	Digital Subscriber Line
UMTS	Universal Mobile Telecommunications System
LAN	Local Area Network
NREN	National REsearch Network
GEANT	Multi-gigabit pan-European data communications network
NTI	Network Technology Independent
NTD	Network Technology Dependent
AQ-SSN	Application Quality Signalling and Service Negotiation
CHAR	CHARging module
QCM	Quality Control Module
MMS	Monitoring and Measurement System
EQ-SAP	EQ-Service Access Point
PQ-WFQ	Priority Queueing - Weighted Fair Queueing
SCTP	Stream Control Transmission Protocol
DCCP	Datagram Congestion Control Protocol
ETP	Enhanced Transport Protocol
gTFRC	TCP-Friendly Rate Congestion Control
TC	Time Constraints
SACK	Selective ACKnowledgement
PCMA	Pulse Code Modulation a-law
PCMU	Pulse Code Modulation mu-law

CIF	Common Intermediate Format
QCIF	Quarter Common Intermediate Format
SQCIF	Sub Quarter Common Intermediate Format
e2e	end-to-end

Chapter 6
The EuQoS System

Michel Diaz, José Enríquez-Gabeiras, Laurent Baresse, Andrzej Beben, Wojciech Burakowski, María Ángeles Callejo-Rodríguez, Jorge Carapinha, Olivier Dugeon, Ernesto Exposito, Mathieu Gineste, Enzo Mingozzi, Edmundo Monteiro, Antonio Pietrabissa, Florin Racaru, Jarosław Śliwiński, Giovanni Stea, Halina Tarasiuk, Nicolas Van Wambeke, and Markus Wulff

Abstract The project "End-to-End Quality of Service support over heterogeneous networks" (EuQoS) is an European research project which has defined a novel architecture that builds, uses and manages the end-to-end (e2e) application exchanges and network paths with Quality of Service (QoS) guarantees across different administrative domains and heterogeneous networks. This chapter presents the architecture of the EuQoS system as a case study of the concepts introduced in previous chapters. The EuQoS architecture provides a clear interface that allows the end user to request a specific QoS level, without changing its application signalling protocol and using the basic connectivity of the local service provider. A complete set of supporting functions has been implemented: i) Security, Authentication, Authorisation and Accounting (SAAA); ii) Admission Control; iii) Charging; iv) Signalling and Service Negotiation; v) Monitoring and Measurements Functions and System (MM-F/MMS); vi) QoS Routing (QoSR); vii) Failure Management; viii) Traffic Engineering and Resource Optimisation (TERO). The EuQoS system has been deployed as a prototype including all the above features, encompassing the most common access networks, i.e., xDSL, UMTS, WiFi, and Ethernet, connected through a core network composed by the National Research and Education Networks (NRENs) of the project partners and GÉANT (the European research network). This section describes the main features of the EuQoS system and presents the mechanisms, algorithms and protocols that have been developed in the project. The results achieved validate the design choices of the EuQoS system, and confirm the potential impact that this project is likely to have in the near future.[1]

[1] This work was partially funded by the European Commission through the EuQoS Integrated Project (contract FP6-004503)

6.1 Introduction

Due to the increasing demand for using new multimedia applications over the Internet (such as VoIP, video streaming or telemedicine), the provision of QoS to these novel services is becoming a key driver for ISPs in the future Internet. In this context, the main challenge is to guarantee the users QoS requirements between the end points involved in the communication; a new architecture is needed in order to address this goal. Its main feature is the integration and synchronisation of the tasks performed in the different planes of the networks along the end-to-end path. In order to address this issue, the EuQoS system has been designed to provide guaranteed e2e QoS over different underlying network technologies. The EuQoS system builds, uses, and monitors e2e QoS paths across different administrative domains in heterogeneous networks.

This chapter presents the final architecture of the EuQoS system as a case study of the concepts introduced in previous chapters, providing a view on how QoS delivery can be supported in real environments using state of the art technologies. The different aspects of the architecture and the implementation of the EuQoS system are introduced in the next sections:

- In Section 6.2 *a top level description* of the architecture and the main characteristics of the EuQoS system are introduced. This high level view presents the key actions and protocols used to coordinate the different technologies and domains in the e2e path. The behaviour of the network and the application levels, together with the way the main system components work, are described.
- Section 6.3 contains *the functional description* of the system based on the three main network design processes, i.e., Provisioning, Invocation and Operation, Administration and Management (OAM).
- Section 6.4 presents *the framework for QoS provision*, which specifies the EuQoS Classes of Service (CoS) and presents how they can be supported in different underlying network technologies.
- Section 6.5 shows, after the signalling and control phases, how the data will be transferred using an adequate transport layer. The six different transport layer services now needed for handling the e2e application-to-application QoS for the different underlying network CoSs are presented.
- Section 6.6 introduces the novel approach selected in the system to implement QoS multicast services. The EuQoS Multicast Middleware uses Scribe & Pastry for defining the Peer-to-Peer (P2P) network and for building the multicast trees. Pastry is a P2P routing substrate and Scribe builds an overlay structure on top of Pastry for multicast tree construction.
- Section 6.7 provides *a real world example* of how commercial applications can be integrated into the EuQoS system. A telemedicine application (Medigraf) is introduced, and the key aspects needed to integrate it in the EuQoS environment are shown.

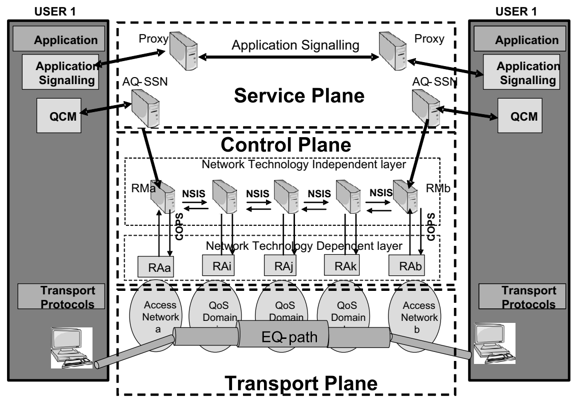

Fig. 6.1 EuQoS end-to-end network architecture

6.2 Architecture

6.2.1 Goals and Requirements

Following the *divide et impera* premise, the system is founded on a division of the e2e QoS paradigm along the vertical axis (Service, Control and Transport Planes) and the horizontal axis (the various network technologies, i.e., the access and the core networks). This is illustrated in Figure 6.1.

Application signalling allows the caller to contact the callee, obtaining its IP address, and to agree on the codecs to be used. It works exactly the same as in the standard Internet nowadays.

The Service Plane, offers access to the EuQoS “QoS on demand” service to provide QoS connections using specific signalling, requesting the necessary resources to the network. Finally, this level is also responsible for authorising, authenticating and accounting of the user activity, and of filtering the QoS requests according to the user profile.

The Control Plane implements the mechanisms to translate the application requests to the network layer, and coordinates the e2e path management. The easy deployment of the EuQoS system has been a key design principle, so that to facilitate different domain providers to adopt the EuQoS solution. This has been met by the specification of a Network Technology Independent level (NTI), responsible for managing the domain at IP level, and a Network Technology Dependent level (NTD), for example performing the algorithms specific for each underlying network technology. The clear interface between them allows any provider to be integrated in the e2e QoS solution by just implementing its own components for the NTD.

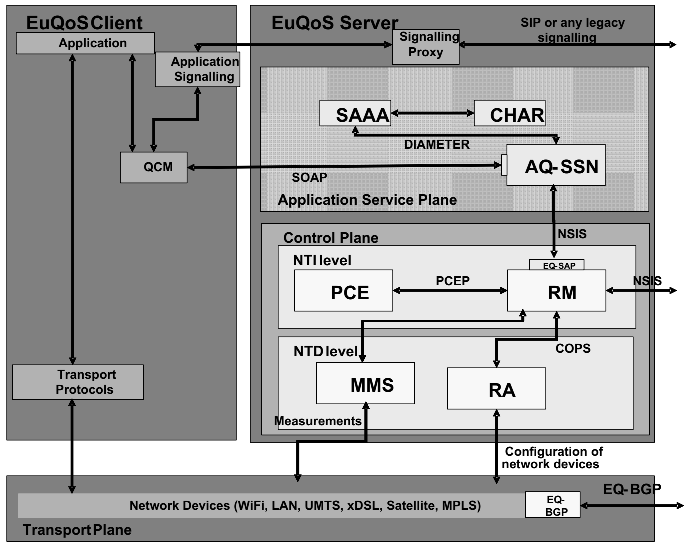

Fig. 6.2 General EuQoS Architecture

The Transport Plane builds the actual e2e paths for the specified Network Classes of Service. It also includes in the hosts a new transport layer protocol which, can be optionally used to provide to the applications different Transport Classes of Service, optimising the data transfer depending on the QoS requested by the applications and the selected Network Class of Service.

6.2.2 Functional Blocks and their Main Functions

Figure 6.2 gives a more detailed view of all these interfaces and of all functional entities composing the EuQoS system, and located at both the client and server sides. As shown in this figure, two sides are well differentiated in the EuQoS system: the EuQoS client and the EuQoS server. At the EuQoS client side, the main functions, located at the user equipment/host, are:

- The **Application** that the customer wants to use.
- The **Application Signalling**: it allows the caller to contact the callee side and to agree on their session parameters, e.g., codecs. This function can be performed by any legacy signalling protocol (as SIP).
- The **Quality Control Module** (QCM) is responsible for managing the data structures as required by the EuQoS server and of asking the EuQoS server as-

sociated to its access domain to establish a QoS session, using the EuQoS "QoS on demand" service.
- The **Transport Protocols** that allow the application to send data to the Transport Plane in the network with an optimizing protocol.

At the EuQoS server side, the structure of the different planes are as follows:

- The **Service Plane**: This plane must allow the EuQoS clients to request the establishment/release/modification of an EuQoS session with e2e QoS guarantees. In this plane, the key function is the Application Quality Service Signalling Negotiation (AQ-SSN) module, which provides the "QoS on Demand" Service to the end user. This plane also supports authorisation, authentication, accounting, and billing for each user session. The SAAA is the module responsible for managing user accesses to network resources (Authentication), to grant services and QoS levels to the requesting users (Authorisation) and to collect accounting data (Accounting), while the CHAR module is responsible for charging the EuQoS customers and managing the bills.
- The **Control Plane** manages the Transport Plane in order to provide the e2e EQ paths (e2e QoS paths), according to the requests received from the Service Plane. So, the Control Plane have to enforce the QoS in its domain underlying technology of its domains and to synchronise this process with the other domains involved in the provisioning of the EQ path. It is split into two different levels:
  - **The Network Technology Independent** level (NTI) is responsible for managing the domains at IP level. This level considers an abstraction of each domain including its topology. The main blocks at this level are the Resource Manager (RM) and the Path Computation Element (PCE).
  - **The Network Technology Dependent** level (NTD) is responsible for performing the resource reservation/release, provisioning of resources, configuring the network elements and algorithms, using the Resource Allocator (RA) element. The Measurement and Monitoring Functions and System (MM-F/MMS) is located at this level.
- The **Transport Plane** composed of the network devices that should be managed by the Control Plane. The main goal of the EuQoS system Transport Plane is to build, use, and manage the EQ paths across all different underlying network technologies.

It is important to note that the interaction with the EuQoS system does not imply the usage of a specific Application Signalling Protocol (such as SIP, H.323, etc.). This allows the easy integration of any application with the EuQoS system: the user must only use the QCM to invoke the "QoS on demand" service in order to request e2e QoS guarantees.

One of the major strengths of the EuQoS system is the clear specification of the interactions between the involved entities: clients and QoS provider, Service, Control and Transport planes, and EuQoS systems located at different ASs. The main interactions in the EuQoS system are:

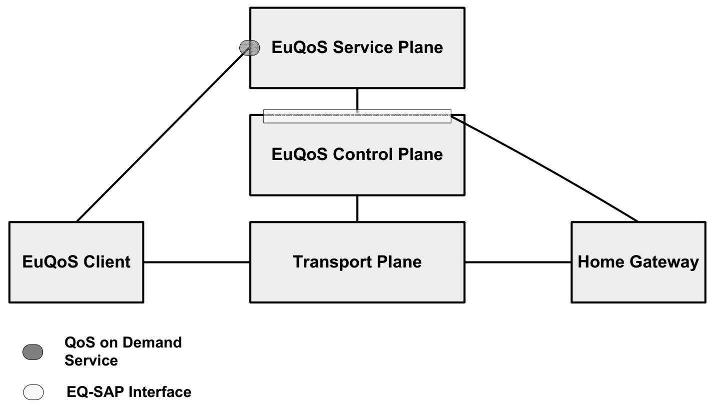

Fig. 6.3 Interaction between the EuQoS Client and the EuQoS server

- Application interaction: It allows the users to contact each other and to agree on the codecs that can be used to start the EuQoS session. Standard SIP is mostly used, but any other legacy application signalling could be used.
- EuQoS Client to server interaction: In order to setup QoS connections from the client side, several approaches can be followed:

  – EuQoS aware applications: This approach considers the application as part of the EuQoS system. In this way the Application invokes the QCM module to provide the QoS connection, and, when an application signalling event is detected, the QCM contacts the AQ-SSN through a Simple Object Access Protocol (SOAP) interface to forward the request.
  – EuQoS non-aware applications: it allows any legacy application to use QoS connections even when it is not integrated in the EuQoS system. To do this, an external program (like a web application) can use the QCM at the client side to ask the AQ-SSN to establish/release/modify EuQoS sessions.
  – Home Gateway integration: When the operator managed equipment (Home Gateway (HG)) represents the boundary between the operator network and the home network, the interaction can be considered to be of an inter-domain type. In this context, the Home Gateway can be considered as an extension of the EuQoS Control Plane, that interacts with its associated Operators Control Plane by means of the EQ-SAP interface to request e2e QoS guarantees in the segment that cannot be managed by the HG. These two interactions between the client and server sides are shown in Figure 6.3.

- Interaction between the Service Plane and the Control Plane: The AQ-SSN module requests the services of the Control Plane using the EuQoS Service Access Point (EQ-SAP) interface, that is implemented using the NSIS protocol .
- Interaction between NTI levels located in different domains: The NSIS protocol is used in order to exchange QoS invocation between different ASs (required to provide e2e QoS).
- Interaction between NTI and NTD levels: COPS primitives are used to ask for the resource reservation and commitment.

In addition to the mentioned functions, the following signalling protocols have been also implemented as part of the EuQoS system:

- Diameter allows the authentication, authorisation and accounting information exchange between the SAAA server and the AQ-SSN module.
- The EQ-BGP routing protocol conveys QoS information between each AS in the global system.
- The Path Computation Element (PCE) Protocol (PCEP) allows the communication between different PCEs of the hard model sub-sets of the EQ path.

6.2.3 Control Plane Elements: RM and RA

As explained above, the management and signalling at the Control Plane is mainly implemented by two components/entities, the Resource Manager (RM) at the Network Technology Independent level and the Resource Allocator (RA) at the Network Technology Dependent level.

6.2.3.1 Resource Manager Architecture

The RM is the Network Technology Independent entity responsible for managing the invocation and provisioning processes (see Section 6.3). RM entities can be deployed in each domain according to the size of the domain.

The RM provides the interface to the Service Plane and to trusted terminals, called EQ-SAP (EQ-Service Access Point), in order to allow these entities to request QoS guarantees for specific flows, while it also provides the interface to the RMs belonging to other network domains involved in provisioning e2e QoS guarantees. The main functions performed by the Resource Manager are:

- RM supports resource and admission control within a single administrative domain and between administrative domains: The RM is the core element of the EuQoS system that contacts the technology specific Resource Allocators (RAs) to enforce the admission control decisions. It further contacts the RMs located in the other domains involved in the EQ path and configures the resources for guaranteeing the QoS requests.

- Verification of resource availability on an e2e basis: The RM applies an e2e Connection Admission Control (CAC) that checks whether there is a provisioned e2e path that meets the QoS requests.
- The final decision point is located at the RM, since it should decide the admission/rejection of a new session according the reservation results in its domain and in other domains.
- Network selection: The RM locates the core networks (via NSIS protocol) and the RAs that enforces the final admission decisions.
- The RM checks whether the connection requests meet the operator policies for this domain. These policies are a simple set of conditions formulated as the maximum bandwidth and QoS parameter limits supported by this domain for each e2e CoS.
- Network topology maintenance: The RM maintains the inter-domain topology used during invocation process.
- Network resource maintenance: The RM maintains information about the expected usage of resources and collects information from different measurement MMF/MMS tools to infer the current usage of the network resources.

There is a complimentary element in the NTI level, called the Path Computation Element (PCE), which is used during the provisioning process in case that MPLS-TE technologies are used (see Section 6.3.1.1). The rationale behind the PCE is to delegate the computation of the best MPLS path to a dedicated server, offloading the RM from this specialised task.

6.2.3.2 Resource Allocator Architecture

The Resource Allocator (RA) is a technology-dependent module responsible for providing and managing QoS in the underlying networks. The RA enforces the traffic handling rules to implement the Classes of Service (CoS) in each network, as specified in Section 6.4. In general, the RA performs the tasks that come from the provisioning and invocation processes and from the monitoring functions (see Section 6.3).

The EuQoS architecture now assumes that a single RA (see Fig. 6.4) is deployed in a given domain and that it manages all the resources that are critical from the point of view of QoS assurance. A pool of RAs could be used instead. The main functionalities covered by this element are the following:

- QoS and priority mapping technology dependent: The CAC makes the final mapping from e2e CoSs (network CoSs) to technology dependent CoSs.
- Gate control: This function is limited and exists only if particular technology operates in a gateway (UMTS, possible for xDSL).
- IP packet marking and rate limiting control: If a given technology is able to perform this function, the RA triggers this feature. Otherwise, one must provide a traffic conditioning module that marks packets generated by end users when they enter the network.

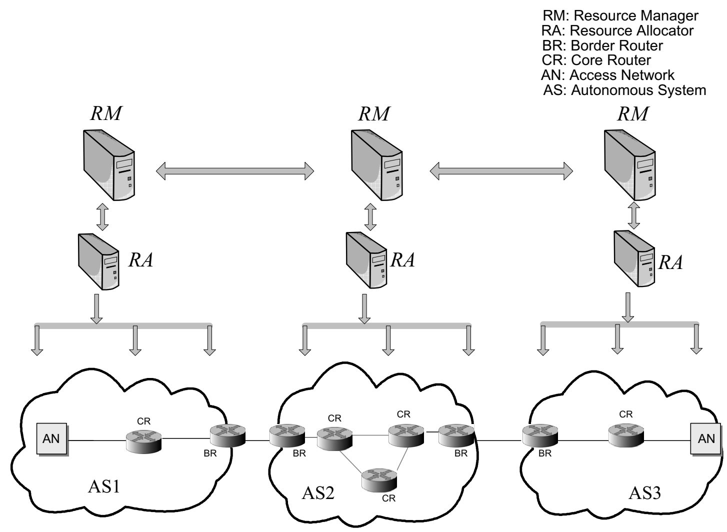

Fig. 6.4 Reference locations of RAs in EuQoS

- Technology dependent decision point: The RA will be responsible of accepting/rejecting one connection request to the specific technology policies.
- Network topology maintenance: The topology information managed by the RA is reduced and covers only access networks operating below IP level. Particularly, when the dynamic IP address allocation is used, the RA must be able to find out the exact location of the user.
- Network resource maintenance: The RA controls resources taking into account provisioning and invocation point of view.
- Element resource control: The RA provides configuration and management of transport elements not only at aggregate level, but also per-flow if access technology allows for it.

6.3 Provisioning, Invocation, and Operation, Administration and Management

EuQoS QoS guaranteed paths (EQ paths) are the EuQoS defined QoS paths providing a given end-to-end QoS. They are implemented in the Transport Plane over a wide variety of technologies and networks, and are built, used, and monitored by the Control Plane in order to provide the QoS needed by the Service Plane. The purpose of these EQ paths is to provide quality guarantees to applications on an e2e basis.

Each EQ path corresponds to a given set of QoS parameters, i.e., those corresponding to the selected Class of Service (CoS). The EuQoS system acts at three different levels:

- The provisioning process is responsible for building the EQ paths across network domains at both independent and dependent network levels. The time scale is in the order of hours or days and it is triggered according to inter-operator agreements.
- The invocation process uses the EQ paths by selecting the most appropriate one, and performs CAC to protect EQ paths from congestion. This process is triggered by the end users when a new session request is sent to the EuQoS system.
- The Operation, Administration And Maintenance (OAM) process protects EQ paths from failure and interacts with the provisioning and invocation processes to repair EQ paths if needed. It also provides the necessary supervision and measurement functions.

In this section, a more detailed description of the provisioning, invocation, and operation, administration and management (OAM) processes is provided.

6.3.1 Provisioning Process

The provisioning process is responsible for:

- computing and setting up e2e data paths between access networks,
- provisioning resources across the different ASs along the path so that QoS guarantees are enforced.

The provisioning process is managed by the Traffic Engineering and Resource Optimisation (TERO) module inside each RM.

6.3.1.1 Resource Provisioning

The EuQoS provisioning process defines two provisioning models, namely the *Loose Model* and the *Hard Model*, which integration allows providers to control the balance between manageability and scalability of the system.

Loose Model

The loose model designs the transport path (between the sending and receiving entities) by starting from the data path. The data path is first selected by a routing protocol, and then the signalling protocol has to reserve the resources for this data path. In the loose model, resources are independently provisioned in every AS. Although resources are provisioned per CoS, there is no specific binding of reserved resources to e2e paths (EQ paths). Therefore, the resources required for establishing a single user connection along an EQ path are dynamically composed and associated to that path by the Call Admission Control (CAC) function at connection setup time, i.e., during the invocation process.

EQ paths are established by means of an EuQoS handling (EQ-BGP) of the Q-BGP protocol, an inter-domain QoS routing protocol, whose objective is to establish e2e paths that offer the most suitable QoS guarantees, taking into account the QoS capabilities of each domain (see section 6.3.1.2). EQ-BGP advertises the reachability of given destinations for each CoS, together with an estimate of the e2e QoS along the selected EQ path.

The main advantage of the loose model is that it requires minimum coupling among the Autonomous Systems (AS) along the EQ path. In fact, it only requires peering agreements between neighbouring Autonomous Systems, without any e2e concept (and related management requirements). As such, it can be considered as the basic Internet-wide model, which is suitable for any policies implemented by a provider with the single constraint of supporting EQ-BGP. For technologies that do not support EQ-BGP or for domains where EQ-BGP is not suitable, the solution is to use EQ-BGP in the RM instead of in the border routers. In this case, the multi-hop classical BGP option is used to link the peering entities.

The main disadvantage of the loose model is the amount of signalling involved in the call setup/teardown process, due to the dynamic binding of resources to the EQ path.

Hard Model

The hard model is based on the concept of an EuQoS defined link, called the EQ link. An EQ link is a configured transport path, having known QoS characteristics between any two nodes in different (non-neighbouring) ASs, and behaving like a virtual inter-domain link interconnecting a pair of neighbouring border routers. As such, it is associated to a specific CoS, not to a session (i.e., it carries traffic aggregates). Resources (bandwidth and buffers) are explicitly reserved for its exclusive use as part of the provisioning process. In practice, an EQ link is established as a DiffServ MPLS-TE tunnel, which may span over multiple domains or ASs (see Chapter 3). Thus, it is semi-static, with resources associated to it, and it can be protected against failures. Based on this concept, an EQ path may be simply built, at provisioning time, by establishing a corresponding EQ link on demand across the Internet between two networks.

EQ link establishment needs specific means for the computation of the AS path along which the EQ link is setup. In fact, in today routers, on-line path computation is done at the head-end Label Switch Router (LSR), but this has some limitations. In particular, in an inter-area and inter-AS context, the head-end routers only have a partial visibility of the topology and cannot compute an e2e path. To solve this issue, a two-steps approach is implemented in the EuQoS system as shown in Fig. 6.5.

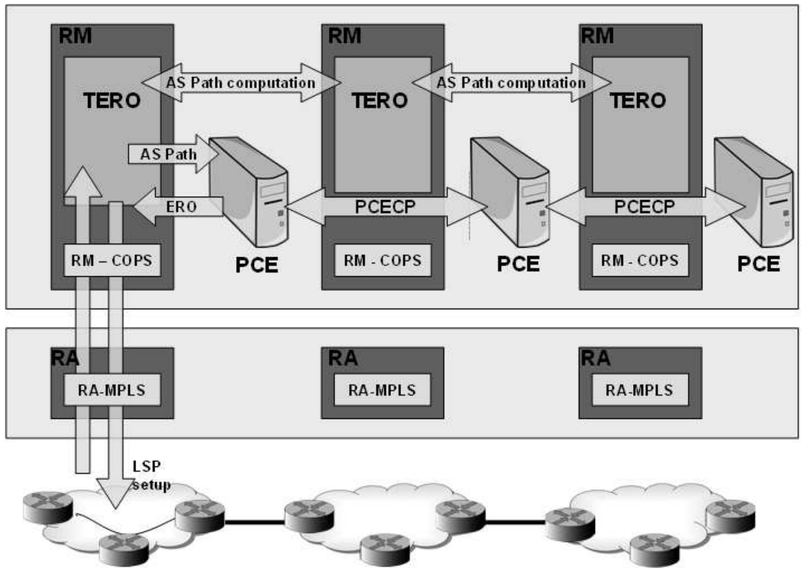

Fig. 6.5 PCE integration in EuQoS

- First, the best AS path between the two ASs is computed through direct interaction of the TERO modules in neighbouring domains. The computation takes into account QoS objectives, resource availability, and administrative constraints that may limit the reachability of the destination with the CoS of the EQ link.
- Then, the actual node-by-node path computation relies on a Path Computation Element (PCE) chain along the computed AS path (see [123]). The rationale behind the PCE is to delegate the computation of the best path to a dedicated server, i.e., the PCE itself. The PCE serves path computation requests sent by a client. Although the original Internet Engineering Task Force (IETF) charter for PCE was meant to take into account only intra-domain path computations, the multi-area was in the scope of the PCE Working Group. In fact, since PCEs can communicate with each other, they can cooperate for computing a path that spans across several ASs. The Path Computation Element Communication Protocol (PCECP) is used for the communication between the PCEs. The result of the computation is delivered as an Explicit Route Object (ERO) to the TERO module.

The EuQoS system relies on the PCE concept to implement multi-domain EQ link setup. A detailed description of the functional requirements and specifications needed to setup EQ links is given in [124].

The advantages and disadvantages of the hard model are opposite to the loose model. First of all, it reduces the signalling required to setup connections. Signalling is only required in the access domains and at the entrance of an EQ link, and resources are already bound at the provisioning time. Furthermore, it can optimise resource provisioning by exploiting inter-domain multi-path capabilities on a per-CoS basis. On the other hand, such a model entails complex path setup procedures, requiring a strong degree of cooperation between remote ASs. Furthermore, it requires support of DiffServ MPLS-TE in the whole core, and as such its applicability is limited to the domains where this mechanism is present.

In summary, regarding the EQ path building there are two possible approaches at the two ends of the spectrum:

- At one end, an EQ path is the result of a sequence (as determined by EQ-BGP) of a number of ingress-egress boarder router paths, each belonging to a single AS, and resources are provisioned e2e per session and per domain as part of the invocation process (loose model).
- At the other end, an EQ path is implemented by a dedicated complete e2e access network-to-access network EQ link, and resources are provisioned e2e per EQ path as part of the provisioning process (hard model). Note that the latter would obviously imply that a full mesh of EQ links connecting access domains could be setup – at least theoretically.

As a a consequence, network provisionning in EuQoS becomes quite flexible, as selecting any combination of loose/BGP-based and hard/MPLS-based paths is possible, depending on different type of constraints, as contexts, agreements, policies, etc.

6.3.1.2 EQ-BGP: Enhanced QoS Border Gateway Protocol

The Enhanced QoS Border Gateway Protocol (EQ-BGP) [125, 126] is the inter-domain QoS routing protocol developed within the EuQoS project. Its objective is to advertise and select the inter-domain routing paths taking into account QoS objectives of e2e CoSs (as defined in Table 6.2). EQ-BGP extends the currently used BGP (BGP-4) [127] inter-domain routing protocol in several ways. First, it defines the QoS Network Layer Reachability Information (QoS NLRI) path attribute that conveys information about e2e CoSs offered on advertised paths. Second, it uses the QoS assembling function for computing aggregated values of QoS parameters guaranteed by each segment of a path. Third, EQ-BGP defines the QoS-aware decision algorithms for selecting routing paths. Fourth, EQ-BGP keeps separate routing table for each e2e CoS.

EQ-BGP performs QoS routing taking into account the QoS guarantees provided by particular domains in multi-domain networks. For that purpose, EQ-BGP routers advertise information about the reachable destinations jointly with aggregated values of the QoS parameters guaranteed by e2e CoSs on currently used paths. Those aggregated values are calculated taking into account the impact of all domains and inter-domain links on the path towards a given destination. Then, the neighbouring EQ-BGP routers update received values of QoS parameters taking into account contribution of their domains and then decide about their routing. In case of any changes, the routers advertise them to neighbours. Finally, EQ-BGP sets the roadmap of paths that are available for all e2e CoSs. The roadmap provides also

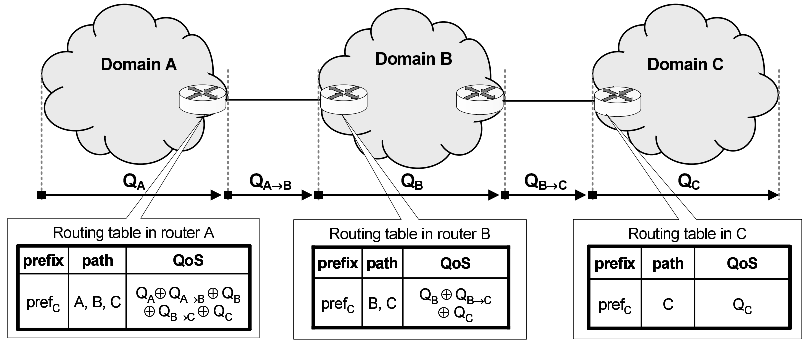

Fig. 6.6 Example of EQ-BGP operation

values of QoS parameters that are guaranteed between each pair of source and destination prefixes.

Figure 6.6 shows an example of how QoS routing information is computed and advertised in the network using EQ-BGP. For the sake of simplicity, we assume a simple network consisting of three domains A, B and C that support only one e2e CoS. Each EQ-BGP router is aware of the values of the QoS parameters that are assured inside its domain (Q_A, Q_B or Q_C depending on the domain) as well as on its corresponding inter-domain link ($Q_{A->B}$ or $Q_{B->A}$, respectively). Those values should correspond to the maximum admissible load that are allowed by the admission control function. The actual values should be fixed during the network provisioning process taking into account details of domain configuration, used technology, provider policies, etc. The values of QoS parameters typically change at provisioning time scales, e.g. in the order of days or weeks, so route changes due to frequent variations of the QoS values are not expected.

Now, let us consider the case when Domain C advertises a new prefix, say $pref_c$. Then, the routing information is propagated towards Domain A through Domain B. Figure 6.6 shows the routing tables of the border EQ-BGP routers along the path. During this process EQ-BGP routers aggregate the values of the QoS parameters taking into account the QoS contribution of particular domains as well as the inter-domain links on the path towards $pref_c$ advertised by Domain C. For example, domain A learns the e2e QoS path towards the destination $pref_c$, with QoS corresponding to $Q_A \oplus Q_{A->B} \oplus Q_B \oplus Q_{B->C} \oplus Q_C$ for considered CoS, wherein the operator $\oplus$ denotes QoS assembling function. Taking into account that QoS parameters used by the e2e CoSs can be treated as additive, we use a simple sum function.

The values of QoS parameters are advertised using the QoS Network Layer Reachability Information (NLRI) path attribute presented in Figure 6.7. The attribute begins with the attribute header that contains flags, type indicator and the attribute length. The flags are used to inform routers that information carried in the QoS NLRI attribute is optional, non-transitive, and complete. The main part of the attribute contains a number of structures describing particular e2e CoSs. Each structure covers the e2e CoS identifier and three fields including IP Packet Transfer

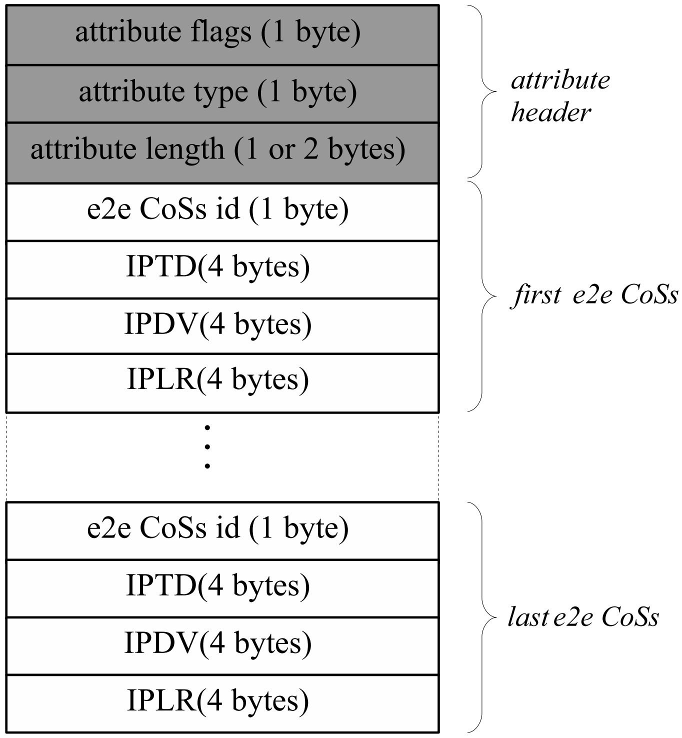

Fig. 6.7 Format of the QOS Network Layer Reachability Information path attribute

Delay (IPTD), IP Packet Delay Variation (IPDV) and IP Packet Loss Ratio (IPLR) parameters. Values of IPTD, IPDV are expressed in μsec, while IPLR is carried in the exponent form: $-1000 * \log_{10}(IPLR)$.

EQ-BGP uses the QoS-aware decision algorithm. It allows the routers to compare the paths going toward a given destination and then to select "the best" one from the viewpoint of QoS objectives of particular e2e CoSs. The algorithm adds a new step in the routing decision process that evaluates the Degree of Preference (DoP) factor based on the values of QoS parameters carried in the QoS NLRI attributes. The degree of preference is used before the path length criterion. So, EQ-BGP will first consider the QoS level offered by the available paths and if this criterion does not decide, the router will select the shortest path. The next decision steps are the same as in case of the BGP-4 protocol.

6.3.2 *Invocation Process*

This section presents the invocation process in the EuQoS system, explaining the signalling chain, the devices and functions triggered in each server involved in a session establishment.

6.3.2.1 Invocation in the Service Plane

The application invocation and signalling phase is used to trigger the application-to-application negotiation and then, if positive, to trigger the network invocation process described in 6.3.2.2.

Taking into account that some applications already have different application signalling, such as H323, SIP or any other ad-hoc protocols, EuQoS proposes a new application level architecture that avoids the restriction of using EuQoS application signalling based on SIP as the only way to interact with the EuQoS network server. The key point of this approach has been to define a "QoS on-demand" service.

Two reference points are being defined to ask for the e2e QoS on-demand service:

- An interface provided by AQ-SSN to the EuQoS clients allows the clients to ask for an e2e QoS request. This interface is implemented using SOAP.
- the RM Service Access Point API (EQ-SAP) for trusted legacy terminals (e.g., as it is proposed in the Home Gateway Initiative) or for any other allowed entity.

It is important to note that this approach makes the clients able to ask directly for QoS parameter reservation to the EuQoS system (that means, the user asks for an e2e CoS for a set of flows, and the user provides its credentials to be authorised and charged) after obtaining the IP address and ports to be used by the callee side.

As explained in the general architecture description (6.2.2), several scenarios can use this architecture. More details can be found in [128].

- EuQoS aware application using QoS-on-Demand service
- QoS-on-Demand services used by administrators for legacy applications, via a web interface.
- Trusted terminals, as home gateways, using EQ-SAP to reserve QoS.

In order to support these scenarios, the main goal of this new approach is to clearly specify the interfaces exposed by the AQ-SSN and RM, QoS on demand and EQ-SAP services, respectively. Table 6.3.2.1 tries to sum up the main characteristics of these reference points.

6.3.2.2 Invocation in the Control Plane

A straightforward invocation process could be as follows:

	QoS on demand service	EQ-SAP
Service provider	AQ-SSN	RM
Service client	QCM and user administrator via web interface	AQ-SSN and trusted terminals
Information exchange requirements	This service must support request-response transactions and shall provide a reliable delivery of the messages.	This service must support request-response transactions and shall provide a reliable delivery of the messages.
Information flows exchanged	Requests: • Perform Reservation • Modify Reservation • Terminate Reservation Responses: • QoS Answer to perform requests. • Result of the reservation termination	Requests: • Perform Reservation Commit • Modify Reservation • Terminate reservation Responses • Resources available to reserve and modify requests • No response to terminate request. It is considered that the connection release is always successful

Table 6.1 Brief description of EuQoS main interfaces

- All domains involved in the EQ path must be asked to reserve the resources corresponding to the connection. This would require a high amount of signalling traffic and a high number of configuration on network equipment.
- The resources in each domain are reserved sequentially. This is not optimal if the setup time is a critical performance parameter, and would have a higher impact if the reservation of all the flows belonging to the same session would be also performed sequentially.

To address the first problem, the hard model has been implemented in the EuQoS system (see section 6.3.1.1). The configuration of transit domains is performed only during the provisioning process instead of during the invocation process. In this way, the signalling load is reduced and the configuration of network equipment in core networks that are supposed to aggregate the traffic from different access networks is not performed during the invocation process. If we consider the simple scenario shown in Fig. 6.8 to explain the invocation chain, the transit domain AS1 will not be asked to reserve resources during invocation process since access networks 1 and 2 will see the EQ path as a link with a specific capacity.

Regarding the second problem, the invocation chain scheme has been designed so that to perform as many actions as possible in parallel.

When the AQ-SSN at the caller side receives the request to establish a new EuQoS session, it can ask the RM to reserve all connections from both directions (caller to callee side or callee to caller) in parallel, without waiting for the first QoS connection request response. The RM will process these requests in parallel triggering all

Fig. 6.8 Invocation scenario

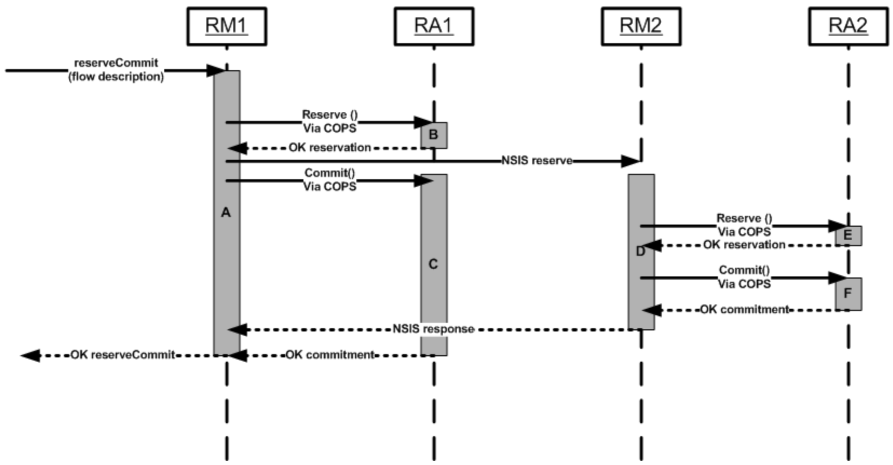

Fig. 6.9 Invocation sequence diagram

the domains involved in the EQ path. The RM located at the caller side will receive all requests needed to reserve the resources for unidirectional flows.

In this scenario two cases can be distinguished, depending on the source IP address of the data flow:

1. The source IP address of the data flow belongs to the (caller) RM administrative domain. In this case the RM receives in the EQ-SAP interface the request to reserve resources for a flow whose IP address belongs to its administrative domain. In order to allow the parallel configuration of network equipment at the access networks, the RM forwards the requests to reserve resources to the next domain after performing the CAC algorithm specific for each technology. The RM effectively reserves the resources while other domains are performing the checking/configuration of their resources. In order to assure that the client has an e2e path with guaranteed QoS, each domain will only send back the confirmation response after receiving the confirmation of the reserved resources from its RA. The sequence of exchanged messages is shown in Fig. 6.9. As it can be seen, this scheme allows configuring in parallel resources in both access net-

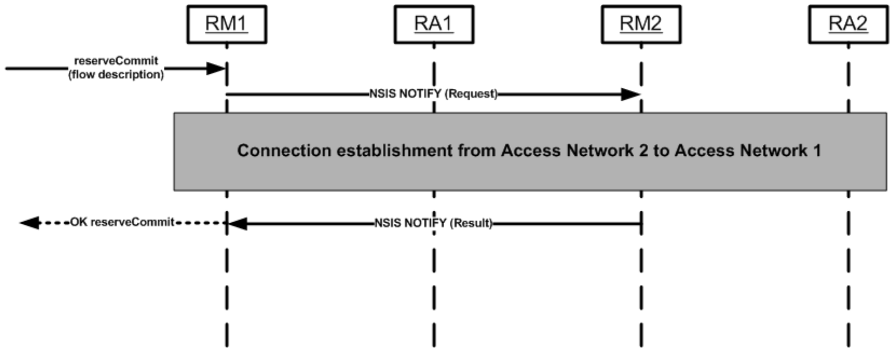

Fig. 6.10 Destination initiated scenario

works. This is interesting, because if, e.g., the first access network is UMTS (the time to establish a session is around 5-10s) and the second is a WiFi domain (this would require around 1-2s), the time required to the configuration of the WiFi equipment would not be added to the time to establish the UMTS session.

2. The source IP address of the data flow does not belong to the RM1 administrative domain. In this case the RM1 (caller side) must resend the request to the RM2 (located at the callee side) and be aware of the result of the reservation. In order to do that, the NSIS NOTIFY message will be used to transport the requests and responses between the access RMs, as shown in Fig. 6.10. The connection establishment from access network 2 to access network 1 follows the description presented in the previous case.

6.3.2.3 Sometimes Per Flow Model

Taking into account the benefits and drawbacks of the loose and hard options mentioned earlier, an intermediate solution has been proposed, named SomeTimes Per Flow (STPF). The details about the STPF model can be found in [129]. The STPF assumes that the resources provisioned for a given CoS in considered domains are divided into two main parts, where one part is reserved only for handling the calls on the basis of the *hard model* scheme (as multi-domains EQ links) while the second part is handled by the *loose model* scheme.

The resources designated to operate loosely per-flow can be used only when there are no resources available in the corresponding hard EQ link. As a consequence, the majority of the call requests should use the hard model, and will not use the full reservation scheme. The full reservation process is then used only for a certain percentage of calls. In this way, it is expected to get high resource utilisation while the required signalling traffic will be noticeably reduced.

6.3.3 Operation, Administration and Management

In order to guarantee the QoS commitment, the EuQoS system performs two actions: the first is the admission control, and the second is the monitoring of the EQ path. This second goal is the main goal of the OAM process. Monitoring is done by means of measurement and fault management.

The measurement sub-system allows the EuQoS system to verify that EQ paths are not overbooked (i.e. the maximum allocated bandwidth corresponds, more or less, to the sum of reserved bandwidth). The fault management sub-system allows verification of the EQ path continuity and takes care of device, node, and link failures. These two sub-systems interact with the invocation process (so that the CAC adjusts the admission control threshold), and the provisioning process (in order to re-compute the EQ path in case of node or link failure). This path protection can be improved by setting up some backup paths by means of a Fast Re-Route (FRR) mechanism when EQ paths are built with MPLS-TE in the hard model.

In order to monitor the provided QoS, the MMF/MMS functions of the EuQoS system monitor the QoS parameters (IPLR, IPTD and IPDV) and the used bandwidth per aggregate. In order to do that, different probes are distributed in each EuQoS domain and the information is reported to all functions involved in the invocation processes. Moreover, the MMF/MMS manages a set of thresholds for QoS parameters and global link utilisation. In case that any of these thresholds is overloaded, an alarm event is generated.

Moreover, for the loose model, the monitoring system will compare the actual EQ-BGP routes with the Service Level Agreement (SLA) information being managed by the TERO module, in order to check that the information agreed between different operators corresponds to the real usage of the network.

The final specification of the functionalities to be covered by the MMF/MMS subsystem of the EuQoS system are described in [130].

6.4 End-to-End Classes of Service in Heterogeneous Networks

This section describes the framework defined in the EuQoS system for providing at the application and at the network layers e2e QoS for heterogeneous multi-domain networks. It presents how connections requiring QoS are established between communicating hosts attached to different access networks. Access networks can be built on different technologies such as xDSL, UMTS, LAN, WiFi, MPLS ad Satellite, and can be interconnected by many IP-based core domains. Furthermore, implementing the framework means to transfer packets while guaranteeing some QoS parameters, i.e. packet delay (IPTD), variation of the packet delay (IPDV) and packet loss ratio (IPLR). The proposed solution should assure that the optimal values of the above parameters are satisfied. The EuQoS approach establishes in the network a number of, so called, Classes of Services (CoSs). The term of Class of Service (CoS) is a service the network offers to traffic streams ([14], [76], [131], [132], [133]).

The rest of this section is organised as follows. Section 6.4.1 describes the implemented e2e CoSs in EuQoS explaining their roles and their QoS objectives. Section 6.4.2 explains the main assumptions that have been made for QoS mechanisms and algorithms required for implementing e2e CoSs in the underlying technologies. It focuses on the specification of generic CAC (Connection Admission Control) algorithms that is the key-element for providing QoS guarantees at the network level. Finally, Section 6.4.3 gives the basic approaches for providing e2e CoSs in each underlying technologies as IP inter-domain links, xDSL, LAN/Ethernet, WiFi, UMTS, MPLS and satellite.

6.4.1 End-to-end Classes of Service in EuQoS

EuQoS assumes that a user can use 6 e2e CoSs (e2e CoSs) that differ in their QoS objectives. A specific CoS is used for handling packets generated by a given type of application as, for example, VoIP connections. Table 6.2 shows the complete set of the CoSs as proposed for the DiffServ architecture [20], [134]). In EuQoS, a subset of these CoSs has been implemented (marked in bold in Table 6.2), as follows:

- The Telephony e2e CoS belongs to the Real Time (RT) class and is mainly dedicated for handling VoIP, emitting streaming traffic of CBR or VBR type. This CoS requires strict QoS guarantees with respect to the selected values of IPTD, IPDV and IPLR.
- The RT Interactive e2e CoS: this class belongs to the RT class and is mainly dedicated for handling VTC (Video-Tele Conferences) as well as interactive games such as NEXUIZ [135] by emitting streaming traffic of CBR or VBR type. This CoS requires strict QoS guarantees with respect to assumed values of IPTD, IPDV and IPLR. This CoS and Telephony CoS differ in packet lengths (rather small for VoIP compared to VTC) and required bandwidth (again, smaller for VoIP) while the required QoS level is similar.
- The Signalling e2e CoS belongs to the RT class and is mainly dedicated for handling application, routing and network signalling traffic. This CoS provides strict guarantees with respect to assumed values of IPTD, IPDV and IPLR. This e2e CoS can guarantee fast connection set-up times. More details about dimensioning this class are in [136].
- The Multi-Media (MM) Streaming e2e CoS belongs to the NRT (Non RT) class and is dedicated for handling streaming traffic (CBR or VBR) generated by VoD (Video on Demand) applications. This e2e CoS provides strict guarantees with respect to assumed values of IPTD and IPLR, but the value of IPDV is not critical.
- The High Throughput Data (HTD) e2e CoS belongs to the NRT class and is dedicated for handling elastic traffic generated by TCP-controlled applications (as in medical applications as Medigraf [137]). As for MM Streaming, this CoS provides strict guarantees with respect to IPTD and IPLR, while the value of IPDV is not critical.

Treatment aggregate	End-To-End Service Class	QoS Objectives			EuQoS Applications							
									Medigraf			
		IPLR	Mean IPTD	IPDV	NEX-UIZ	VoIP	VTC	VoD	VTC	Collabo-ration	Data transfer	Chat
CTRL	Network Control	10^{-3}	100 ms	50 ms								
Real Time	**Telephony**	10^{-3}	100/350 ms (local/long distance)	50 ms		X						
	Signalling	10^{-3}	100 ms	U								
	MM Con-ferencing	10^{-3}	100 ms	50 ms								
	RT Interactive	10^{-3}	100/350 ms (local/long distance)	50 ms	X		X		X			
Non-Real Time / Assured Elastic	Broadcast video	10^{-3}	100 ms	50 ms								
	MM Streaming	10^{-3}	1 s non critical	U				X				
	Low Latency Data	10^{-3}	400 ms	U								
	OAM	10^{-3}	400 ms	U								
	High Throughput Data	10^{-3}	1 s non critical	U							X	
Elastic	**Standard**	U	U	U								X
	LowPriority Data	U	U	U								

Table 6.2 Mapping of EuQoS Applications to Classes of Service

e2e CoS	DSCP Name	DSCP Value
Telephony	EF	101110
Signalling	CS5	101000
RT Interactive	CS4	100000
MM Streaming	AF3x	011xx0*
High Throughput Data	AF1x	001xx0*
Standard	DF	000000

Table 6.3 DSCP codes/names for e2e CoSs in EuQoS (* xx $\in \{01, 10, 11\}$)

- The Standard e2e CoS provides best effort and it means that no guarantee is provided for the IPTD, IPDV and IPLR parameters but the network allocates a given amount of bandwidth to this CoS.

The network will recognise that an IP packet belongs to a given e2e CoS by analysing the DSCP (Differentiated Services Code Point) field in IPv4 or the Type of Service (TOS) field in IPv6. The appropriate code in the packet is assigned by the user equipment and again by the first network element that handles the packet. Table 6.3 shows the DSCP codes/names corresponding to the e2e CoS in EuQoS as proposed in [20].

Figure 6.11 shows the concepts followed for implementing the above specified set of CoSs, regarded as globally known by the users (and the user QoS-aware applications). A user who wants to use a given application (VoD, VoIP etc.) activates its QoS and submits its QoS request to the predefined e2e CoS, accordingly to the mapping given in Table 6.2. In EuQoS, possible paths are EQ paths, and when the path

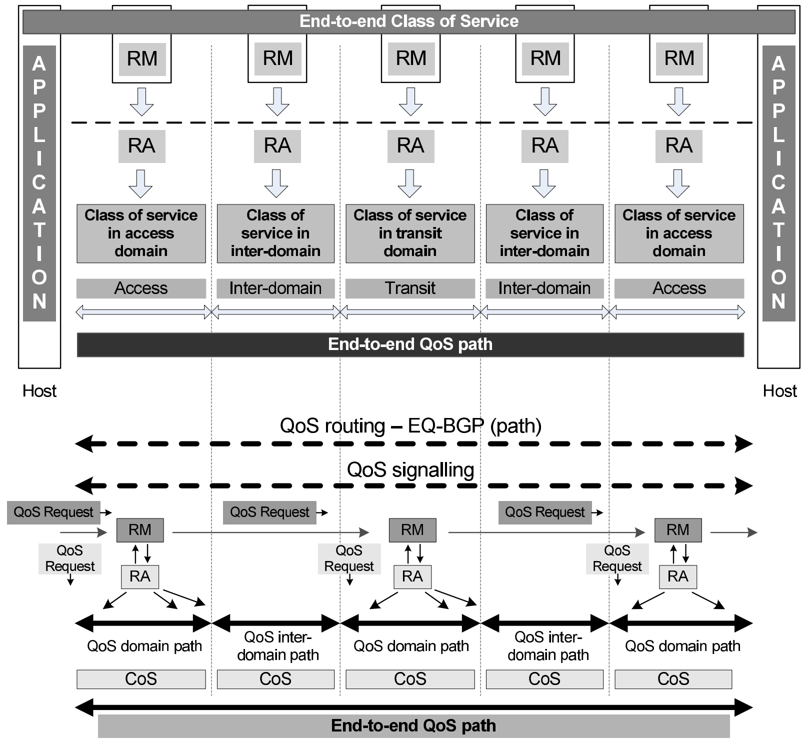

Fig. 6.11 Concept of e2e CoSs for implementation in EuQoS system

is established, the QoS request in sent to the RMs situated along this path. When a RM receives the QoS request, it communicates with its associated RA elements for checking whether the requested resources are available in the underlying network (see section 6.3.2).

The simplest solution occurs when a given underlying technology supports by itself the same CoSs as EuQoS, in terms of handled traffic profiles and QoS guarantees. However, for some underlying network technologies there are not clearly specified CoSs that are compatible with the e2e EuQoS CoSs. So, new EuQoS specific solutions have been investigated and implemented for providing packet transfer capabilities as requested by e2e EuQoS CoSs.

Depending on the capabilities of the network technologies, the proposed solutions are mainly based on providing an adequate Connection Admission Control (CAC) function to limit the QoS traffics, and on tuning the available QoS mechanisms (schedulers, shapers, policers etc.) in the network elements (IP routers, access points in WiFi, LAN/Ethernet switches etc.).

6.4.2 QoS Mechanisms and Algorithms for Specification of e2e Classes of Service

The i-th ($i = 1,\ldots,6$) e2e EuQoS CoS is designed for handling streams having a given traffic profile, i.e., to assure adequate packet transfer characteristics (maximum allowed values for $IPTD_{e2e,i}$, $IPDV_{e2e,i}$ and $IPLR_{e2e,i}$). Furthermore, the i-th e2e CoS over heterogeneous multi-domain network needs a compatible $CoS_{j,i}$ for each domain j along the e2e path ($j = 1,\ldots,N$; N is the number of different domains along the path), also expressed by the above mentioned three parameters, $IPTD_{j,i}$, $IPDV_{j,i}$ and $IPLR_{j,i}$. Due to the additive properties of IPTD, IPDV and IPLR[2], for each i-th e2e CoS we have :

$$IPTD_{e2e,i} = \sum_{j=1}^{N} IPTD_{j,i}$$

$$IPDV_{e2e,i} = \sum_{j=1}^{N} IPDV_{j,i}$$

$$IPLR_{e2e,i} \cong \sum_{j=1}^{N} IPLR_{j,i} \tag{6.1}$$

Note, that in equations 6.1 for a given e2e CoS we take into account only the parameters that are specified.

The general principles used to design CoSs mean: (1) to allocate resources for the considered class, (2) to apply QoS mechanisms (in network devices) for forcing required packet transfer characteristics, and (3) to limit the traffic submitted to these resources by an appropriate CAC.

Let us illustrate these rules by considering an e2e CoS that handles traffic streams described by a Peak Rate (PR) and requiring transfer characteristics not larger than the predefined values $IPTD_{e2e}$, $IPDV_{e2e}$ and $IPLR_{e2e}$. Let us also assume that after the provisioning process, the requirements for a given domain are the predefined maximum values of parameters IPTD, IPDV and IPLR.

Example: Designing CoS with predefined maximum values of parameters IPTD, IPDV and IPLR. The CoS handles the traffic streams with declared PRs.

(1) *Allocation of resources*

The required resources for the CoS are usually represented by the link capacity (C) and an associated buffer (B). The CoS is designed for handling packet streams emitted by applications with similar traffic characteristics. So, for the sake of

[2] For two domains with $IPLR_1$ in domain 1 and $IPLR_2$ in domain 2, the resulting $IPLR$ is $IPLR_{1+2} = IPLR_1 + IPLR_2 - IPLR_1 * IPLR_2$. In practical cases, $IPLR_{1+2}$ is around $IPLR_1 + IPLR_2$ as $IPLR_1 * IPLR_2 << IPLR_1 + IPLR_2$. Therefore it can be considered as additive.

simplicity, we can assume that the applications generate the packets with constant length (L). In this case, we can control IPDV by setting B and C, since:

$$IPDV = \frac{LB}{C} \tag{6.2}$$

Furthermore, the commonly known condition in the case when a number of packet streams is multiplexed on a single link is that the link utilisation should be less than 1. The condition for maximum link utilisation, say ρ_{max}, comes from constraints on IPLR or IPTD. The relations for IPLR and IPTD, derived from the analysis of the M/D/1/B [138] and M/D/1 (e.g., [139]) respectively, are:

$$\rho_{IPTD} = \frac{2B}{2B - \ln(IPLR)} \tag{6.3}$$

$$\rho_{IPTD} = \frac{2(IPTD - T_{prop} - \frac{L}{C})}{2IPTD - 2T_{prop} - \frac{L}{C}} \tag{6.4}$$

where T_{prop} represents propagation delay. Finally, we calculate ρ_{max} from:

$$\rho_{max} = \min[\rho_{IPTD}, \rho_{IPTD}] \tag{6.5}$$

The term 6.3 dominates in the most practical cases and 6.4 occurs only when the links have large propagation delays and rather low capacity C, e.g. for a case where $T_{prop} = 90ms$, $C < 4.4\,MBps$, $B = 10\,packets$, $IPLR = 10^{-3}$, $L = 150\,Bytes$ and $IPTD = 100ms$.

(2) *To apply available QoS mechanisms in devices for forcing required packet transfer characteristics*
The set of QoS mechanisms that are available in network devices differs depending on the underlying technology. Anyway, at least for now, the reference QoS mechanisms are specified as PHB mechanisms in the DiffServ architecture. Assuring the requested packet transfer characteristics is based on the type of available schedulers. The preferred schedulers are Weighted Fair Queueing (WFQ) and Priority Queuing - Weighted Fair Queuing (PQ-WFQ) because they assure isolation between CoSs, i.e., guaranteeing isolated buffer size and a given percentage of the total link capacity. So, the traffic belonging to a given CoS is gathered in a dedicated queue.

(3) *Limiting the traffic submitted*
Limiting the traffic submitted to a given CoS can be obtained by applying the following well known formula for peak rate allocation [138]

$$PR_{new} + \sum_{i=1}^{K} PR_i \leq \rho_{max} C \tag{6.6}$$

e2e Class of Service	Inter-domain Class of Service
Signalling	Signalling (S-CoS)
Telephony Real Time Interactive	Real Time (RT CoS)
Multi-Media Streaming High Throughput Data (HTD)	Non Real Time (NRT CoS)
Standard	Standard (STD CoS)

Table 6.4 Mapping between EuQoS e2e CoSs and the inter-domain CoSs

where PR_{new} is the peak rate of new connection requests while K is the number of running connections, each of PR_i ($i = 1, \ldots, K$). The CAC function is invoked during the invocation process in its setup procedure.

6.4.3 Implementation of e2e Classes of Service in Underlying Technologies

This subsection provides a very brief description of the technology specific CoSs, associated to e2e CoSs as specified by EuQoS. The approaches have been implemented and tested in the PAN-European testbed environment ([124]).

6.4.3.1 Inter-Domain Links

Inter-domain links connect two peering ASs and have two unidirectional links, one for each direction. More precisely, the inter-domain link for one direction begins at the output port at the egress Border Router (BR) in one domain and it terminates at the ingress BR of the peering domain. The Per Hop Behavior (PHB) mechanisms that are implemented in the egress BR, including such schedulers as PQ-WFQ or/and WFQ, can be used.

EuQoS defined four inter-domain CoSs that are: (1) Signalling (S-CoS), (2) Real Time (RT CoS), (3) Non Real Time (NRT CoS), and (4) Standard (STD). Table 6.4 shows the mapping of EuQoS e2e CoSs (see Table 6.2) to the inter-domain CoSs.

For inter-domain, the CAC function is performed in the egress BR, at its output port. Each inter-domain BR follows the DiffServ concept, i.e. packets belonging to Telephony and Real Time interactive streams are treated by the router according to the same PHB as specified for the Real Time CoS, and packets belonging to Multi-Media streaming and High Throughput Data (HTD) are treated by the router according to the PHB defined for the NRT CoS. The details of the system analysis are in [140], [141].

6.4.3.2 xDSL

In Digital Subscriber Lines (xDSL) networks four possible network points are candidates to be the bottlenecks and need to be considered: the user xDSL modem (the gateway/Customer Premises Equipment CPE), the Digital Subscriber Line Access Multiplexer (DSLAM) Aggregation Module, aggregation switch(es) and IP edge node. However, in practice some simplifications can be made, depending on the specific characteristics of the network technologies and the capabilities of particular elements.

It must be clearly stated that the evolution of DSL technology results today in a range of DSL standards (ADSL, ADSL2+, SHDSL, VDSL2, etc.) with different bit rates and architecture, affecting its major building block, DSLAM and Broadband Remote Access Server (BRAS). The market demands for cost-effective, differentiated multimedia services provided in DSL networks. This forces the most popular purely Asynchronous Transfer Mode (ATM) DSLAMs to be migrated to fully IP-aware appliances with Ethernet uplinks in the aggregation segment. This makes DSL architecture more flexible and scalable. For instance, for distributed and small groups of subscribers IP-DSLAM may include the functionality of BRAS in a one equipment. Considering the access part of DSL, one may find customer equipment, which is very simple and limited in functionality, devices without QoS mechanisms as well as fully configurable, DiffServ supporting, manageable gateways, mostly deployed for business customers.

In order to achieve CAC for any variant of DSL access network, the CAC algorithm proposed in section 6.4.2 should be used for every IP-aware port with implemented QoS mechanisms. EuQoS considered the access and aggregation segments and focused on two network elements, the DSLAM (more precisely, the IP DSLAM, to implement the QoS mechanisms for IP traffic) and the IP edge node (BRAS). The proposed CAC algorithms for the above elements differ in their assumed type of CoSs provision. In the aggregation segment, we can apply a static partitioning of the link capacity between CoSs, as e.g. in the inter-domain links, while for the access segment we need to focus on link capacity sharing.

6.4.3.3 LAN/Ethernet

In switched Ethernet, the basic mechanism to differentiate traffic is priority scheduling. According to IEEE 802.1Q [142] and 802.1p (part of the IEEE 802.1D [143]) standards, the MAC layer has specified eight priority levels, each for a different Ethernet CoS. The priority level of a Ethernet frame is marked in the 3 bit priority field. It is important to remark that eight priority levels are not available in all devices and one can find equipment with four or even two priority levels. Table 6.5 shows the proposal for mapping the e2e CoSs into Ethernet CoSs in the case where four priority levels are available.

The implementation of e2e CoSs in LAN/Ethernet is not trivial because of the organisation of buffer management based on a shared buffer architecture. The packets

e2e Class of Service	Ethernet CoS	802.1p: values in priority field in Ethernet frame header
Signalling	Network Management	7 (highest)
Telephony, RT Interactive	Voice	6
	Video	5
MM Streaming, High Throughput Data	Controlled Load	4
	Excellent Effort	3
Standard	Best Effort	0
	Undefined	2
	Background	1

Table 6.5 Mapping between e2e CoSs and Ethernet CoSs

belonging to different CoSs share common buffer space. This space is for all output ports. For providing isolation between CoSs and to control IPLR, it is proposed to explore the following additional features of an Ethernet switch:

- The ability to identify traffic flows based on information at Layer 3 and 4, namely source and destination IP addresses, ports and transport protocol (for EuQoS flow identification) [144];
- The ability to perform data bit rate control on a per flow basis [145], [146];
- The ability to perform random early packet discarding based on the queue size at the Ethernet output port (Weighted Random Early Detection (WRED) mechanism).

The CAC function is performed in two elements of the LAN/Ethernet access networks, in the Ethernet Switch (ES) output port and in the Edge Router (ER). The applied CAC algorithm follows equation 6.6.

6.4.3.4 WiFi

The EuQoS approach for providing e2e CoSs in WiFi technology is based on WiFi Multi-Media (WMM) extension [147] and exploits the Enhanced Distributed Coordination Access (EDCA) protocol defined in the extension. The EDCA protocol allows for differentiation of traffic using 4, so called, Access Categories (AC). However, the EDCA itself does not provide strict QoS guarantees as required for e2e CoSs. Then, our CoSs for WiFi use enhanced ACs with additional QoS mechanisms for: (1) provisioning of network resources dedicated for particular CoSs such as values of bandwidth, buffer size and parameters of the MAC protocol, (2) performing CAC , (3) conditioning the traffic generated by users (packets policing/shaping and marking), and (4) providing packet scheduling at the IP layer in access point (AP).

Table 6.6 shows the mapping between e2e CoSs and WiFi CoSs. The WiFi CoSs real time (RT), non-real time (NRT), signalling (SIG) and best effort (BE) are similar to the ones assumed for inter-domain links (see Table 6.4).

e2e Class of Service	WiFi CoS (WMM AC)	QoS objectives[a]		
		IPTD [ms]	IPDV [ms]	IPLR
Telephony RT Interactive	Real Time (AC_VO)	5	15	10^{-4}
MM Streaming High Throughput Data	Non Real Time (AC_VI)	10	-	10^{-4}
Signalling	Signalling (SIG)(AC_VI)	10	-	10^{-4}
Standard	Best Effort (AC_BE)	-	-	-

[a] exemplary target values assumed in provisioning process

Table 6.6 Mapping between e2e CoSs and WiFi CoSs

The solution for WiFi WMM assumes that a single AP will handle traffic belonging to all WiFi CoSs (including best effort traffic), and the EDCA algorithm allows to provide traffic separation between the CoSs.

6.4.3.5 UMTS

For UMTS, the main recognised problem is due to the lack of open interfaces for controlling the specific QoS mechanisms. As a consequence, for EuQoS it has been decided to look at UMTS from two perspectives: a) UMTS as a black box where available UMTS services are reused b) using an implicit, measurement-based, cell-load control approach that can be achieved by using traffic shaping for the connections with assigned low priority (non-EuQoS connections submitted to background CoS). This second approach also addresses the problem of defending the already established EuQoS connections against QoS starvation. This problem is typical of UMTS networks due to the frequent changes of radio channel conditions.

The first approach has been called Usage of built-in CAC from UMTS and the latter Measurement based Open GPRS Gateway Support Node (OpenGGSN) CAC.

Usage of built-in CAC from UMTS

The main goal of the proposal is to take advantage of the built-in CAC from UMTS, which enables decision-taking based on cell load conditions (different among cells), and seamless resource reservation.

E2e CoSs are mapped to their corresponding UMTS traffic classes. Table 6.7 shows the proposed mapping, taking into account the desirable solution and the availability of commercial equipment.

EuQoS e2e CoS	Ideal mapping to UMTS	Feasible mapping
Telephony	Conversational	Interactive (THP = 1)
Signalling	Background	Background
Real Time Interactive	Streaming	Interactive (THP = 2)
Multi-Media Streaming	Interactive (THP=1)	Interactive (THP=3)
High Throughput Data	Interactive (THP=2)	Interactive (THP=3)
Standard	Background	Background

Table 6.7 Mapping between EuQoS and UMTS CoSs (THP- Traffic Handling Priority)

Measurement based OpenGGSN CAC

Although currently available UMTS deployments provide CoSs with strict QoS guarantees with respect to the values of parameters IPTD and IPDV, such features can not be fully exploited mainly due to some limitations in current operating systems of computer Terminal Equipment (TE) that are connected to UMTS by mobile phones (Mobile Terminal MT).

The problem to solve is the preservation of already established sessions with a guaranteed quality in presence of dynamic changes of the radio channel. In this case, even for the admitted connections we need additional mechanisms as continuous monitoring and prioritised treatment in order to maintain the assumed QoS level for them.

One feasible approach to prevent unexpected reconfiguration of resources is to reduce transmission rates to accommodate low priority (non-EuQoS) users with worse radio channel characteristics. Anyway, non-EuQoS users with better radio channel characteristics may maintain connections with unchanged bit rates if there are available resources. Some additional architecture components have been developed and deployed (protocol analyser with online session tracing and logging) to implement this function. To do this, the OpenGGSN basic functionality was improved by adding standards compliant secondary Policy Decision Point (PDP) context management.

6.4.3.6 MPLS (DiffServ-TE)

The QoS enforcement in the hard model is based on a two step approach. The first one consists of provisioning and reserving bandwidth for an EQ link, i.e. an Label Switch Path (LSP) of a given CoS. The second one consists of preventing an excessive amount of traffic to be routed through an LSP, which is accomplished by performing the usual CAC at the EQ link head-end before accepting the new session.

During the EQ link setup the bandwidth is guaranteed as follows:

- RSVP-TE reserves logical bandwidth for a given CoS. The remaining bandwidth for the CoS is automatically advertised by the TE-routing protocol
- Each LSP of the same CoS shares the same queue, buffer and scheduler.

e2e Class of Service	DiffServ MPLS Class of Service
Signalling	CS5
Telephony Real Time Interactive	EF
Multi-Media Streaming High Throughput Data	AF
Standard	BE

Table 6.8 Mapping between EuQoS e2e CoSs and DiffServ MPLS CoSs

The PCE server compares the requirements of a tunnel against the remaining bandwidth of the CoS pool at each router as it performs the provisioning CAC. The bandwidth is maintained by the TE-routing protocol to protect the resource pool against overbooking.

On a link, four classes will have a guaranteed bandwidth allocation: Signalling, Real-Time, Non Real-Time, and Class Default.

So the QoS guarantee first provisions and reserves bandwidth for LSPs in a given CoS, and second protects the LSP against too many flows by performing usual CAC before accepting a new session.

6.4.3.7 Satellite

Scheme for Assuring QoS

The Satellite System provides an access network using the Digital Video Broadcasting - Satellite (DVB-S) and the Digital Video Broadcasting - Reverse Channel Satellite (DVB-RCS) standards to carry out IP-based applications over geostationary satellite. The main concern in the satellite communication is to make an efficient use of the scarce and costly resources. The asymmetric nature of the satellite communication architecture involves different mechanisms to manage resource access.

Static and dynamic access techniques for satellites have been designed and integrated into the Demand Assignment Multiple Access protocol (DAMA) for the DVB-RCS standard, in order to ensure a high utilization of the return link resources and offer QoS-oriented capacity assignment. DAMA access supports four main capacity assignment types to reach its objective:

- Continuous Rate Assignment (CRA): Static and fully guaranteed rate capacity.
- Rate Based Dynamic Capacity (RBDC): Guaranteed capacity up to $RBDC_{max}$ ceiling rate, but this requires dynamic requests (on-demand capacity).
- Volume Based Dynamic Capacity (VBDC): The capacity is assigned when available in response to a request without any guaranty on assignment.
- Free Capacity Assignment (FCA): Automatic allocation of unused capacity, no guarantee and no requests are associated with this assignment type. Because of this automatic allocation, FCA type is not used in the EuQoS services.

DAMA Classes	EuQoS e2e Class of Service					
	Telephony	RT Interactive	Signalling	MM Streaming	High Throughput Data	Standard
CRA	**X**	**X**	**X**			
RBDC				**X**	**X**	
VBDC						**X**

Table 6.9 Mapping from EuQoS CoS to DVB-RCS access classes

Thus, the satellite lower layers are able to provide different types of service, while keeping efficient link resources utilization.

The CAC algorithm performed by the satellite RA in EuQoS benefits from the DAMA access scheme, but also from the information provided by the Network Control Centre (NCC) concerning the agreement passed between the satellite terminal and the satellite system.

Table 6.9 summarises the mapping between RT, NRT and Standard CoSs and the DVB-RCS access classes.

6.5 EuQoS Enhanced Transport Protocol

6.5.1 Introduction

Past and new generations of transport layer protocols have been designed taking into account only a subset of the requirements of multimedia applications. These requirements are basically characterised by reliability and order constraints. Indeed, existing protocols have been designed to provide full order and full reliability (i.e. TCP and SCTP) or no order and no reliability at all (UDP and DCCP). Even if DCCP estimates network congestion by detecting packets out of order, it does not implement any mechanism to deliver packets in any particular order.

At the network layer, standard (Best-Effort) service is still the predominant network service in the Internet, but new network services are proposed, as in EuQoS. Additionally, emerging wireless, mobile or satellite technologies present different network characteristics that should be considered by transport protocol designers, which means for instance to handle variable delay and packet loss rates induced by physical channels.

All these reasons led us to propose an EuQoS Enhanced QoS-oriented transport protocol, here noted as EQ-ETP, intended to provide optimised and differentiated e2e transport layer services for multimedia applications using the different available network layer CoSs.

Mechanisms implementing these transport layer services have to be designed such that they can respond to the various application requirements using the services provided by underlying heterogeneous networks. Moreover, an Enhanced Transport

Protocol (ETP) should be designed within an extensible framework aimed at integrating future mechanisms intended to satisfy new requirements and/or to operate under new networks.

The EuQoS Enhanced Transport Protocol is presented in detail in [124]. In the following, a brief overview of the different service compositions for the EuQoS network services is given.

6.5.2 Enhanced Transport Protocol Services for EuQoS

Enhanced Transport Protocols aim at fulfilling the multiple QoS requirements of multimedia applications over best-effort networks. EQ-ETP extends ETP for handling multiple CoSs. As the corresponding protocols and solutions need to be deployed over different network services, they should be implemented using a dynamic architecture. A flexible and compositional architecture has been designed and implemented in ETP in order to achieve a polymorphic deployment of various internal mechanisms suited to manage the multiple QoS requirements of applications over the various classes of services provided by EuQoS. This architecture allows QoS control and management mechanisms to be easily deployed and configured in order to efficiently work together.

The modular approach of ETP has been defined in order to provide an effective way to satisfy a large range of applicative requirements by adequately composing and fine-tuning different well identified and designed transport layer building blocks (rate control, shaping, congestion control, flow control ...).

Given the nature of the EuQoS network classes of services, various possible compositions have been developed for EQ-ETP in order to provide the most adequate transport layer services regarding the temporal requirements of both streaming applications and non-streaming applications. These compositions are presented in Table 6.10 where RC is Rate Control, EC is Error Control, TFRC and guaranteed TFRC (gTFRC) are congestion control mechanisms.

	Streaming ***e.g. VoD***	**Non-Streaming** ***e.g. file transfer***
Telephony - RT Interactive	**ETP [RC]**	**ETP [RC + EC]**
MM Streaming - HTD	**ETP [gTFRC]**	**ETP [gTFRC + EC]**
Standard	**ETP [TFRC + TC]**	**ETP [TFRC + EC]**

Table 6.10 EQ-ETP Service Composition for the EuQoS traffic classes

6.5.3 Services for Streaming/Non-Streaming Applications

In the following, the different EQ-ETP compositions for each of the EuQoS traffic classes are presented. Two possible combinations of these are possible depending on the nature of the applications. Streaming applications, which transmit one or more multimedia flows, have specific requirements concerning error and time control while non-streaming applications (generally speaking FTP like applications) require full reliability as their time constraints are of less importance than the ones of streaming applications.

6.5.3.1 Real Time Classes of Service (Telephony, RT Interactive)

In case of the Real Time Classes of Service, the application respects the traffic profile it has issued a reservation request for, reliability is guaranteed for the whole stream throughout the EuQoS system. In this context, as streaming applications are generally able to specify their bandwidth requirements accurately, the service composition is limited to a dynamic binding to the UDP protocol. Optionally, the operators might specify that a traffic shaper must be instantiated for shaping at the sending host. Thus, the load on the system routers is reduced. In this context, non-streaming applications generally have no loss tolerance. As the application might have underestimated its resource requirements, the transport services are composed of a shaper coupled to a SACK based error control to provide full reliability.

6.5.3.2 Non-Real Time Classes of Service (MM Streaming, HTD)

In case of Non Real Time Classes of Service, the application might exceed the traffic allowance that it has issued a resource reservation for. In such scenarios, the excess traffic competes with other flows for which it has to respect certain friendliness in order to avoid network collapse caused by congestion. This is achieved by means of the gTFRC module [148] as described in Section 5.4.

In the case of a non-streaming application, the zero loss tolerance is tackled by the addition of a SACK based error control mechanism to ensure the correct, ordered delivery of packets.

6.5.3.3 Standard Class of Service (Best Effort)

In the Standard Class of Service, all traffic must be shaped according to a congestion control algorithm in order to protect the network against congestion collapse. In order to improve the QoS provided to multimedia streams, a TC (Time Constraints) module will be used to offer fast retransmission mechanisms when the time dependence of the packets (VoD Scenarios) [149]. In these scenarios, as non-streaming applications have total reliability requirements, a SACK based error control is added to the composition. Furthermore, as time constraints are proper to streaming applications, the TC module is not enabled for non-streaming applications.

6.6 Multicast

If a data packet should be sent to more than one destination, the sender usually sends the same packet as many times as there are receivers interested in getting the data. Therefore, multiple point-to-point connections are established. This one-to-one communication paradigm is called unicast . In the early days of the Internet, when email, FTP and remote host access were the main applications, there was no need for other paradigms. But the Internet has changed a lot since then. Particularly, the appearance of the Web changed the situation. Now, pictures, movies and audio/video streams are available over the network and their transmission uses up a significant portion of the available bandwidth. With today's technology it is possible to afford a unicast connection for everyone who wants to view a web page. However, to send live audio and video data, which needs a huge amount of network resources compared with web pages, it is not reasonable to have a single connection to each receiver.

The drawbacks of the unicast approach for this kind of applications are evident. First, the source is required to hold a complete list of receivers and second, multiple identical copies of the same data flow over the same links. Instead, data to multiple destinations can be delivered using multicast [150]. Multicast allows the source to send a single copy of data, using a single address for an entire group of receivers. Routers between the source and receivers use the group address to route the data. The routers forward and duplicate data packets wherever a path to receivers diverges.

IP knows three basic addressing modes. A unicast packet is sent to one receiver, a broadcast packet is sent to all hosts of a subnet and a multicast packet is directed to a group of receivers. Unicast and broadcast can be seen as multicast communication with the group of receivers containing one or all hosts respectively. A fourth mode is called anycast. This is a routing scheme which delivers the packet to the "nearest" (considering an appropriate metric) host out of a group of receivers.

The only difference between unicast and multicast addressing from the IP layer's point of view is the usage of special IP multicast addresses [151]. Unlike the unicast addresses a multicast address is not assigned to a single host or network interface. The 32 bit address space of IPv4 (IP version 4) has been divided into five address classes A, B, C, D and E. The most significant bits of an address define its class. The multicast address class is sometimes referred to as class D. Unlike the address classes A, B, and C the multicast address has no further structure. In case of the new IP version 6 (IPv6), all multicast addresses begin with the format prefix FF_{16} [152].

The Internet group management protocol (IGMP) [153] allows the hosts in the Internet to join and leave multicast groups. In order to reduce the amount of data sent over the network links, IGMP manages dynamic groups of multicast receivers.

The group management is done by the routers. Therefore, every router remembers the hosts connected to its local interface(s), which are interested in receiving multicast data and the respective multicast group IDs. IGMP provides the functionality for hosts to tell the routers in which multicast groups they are interested. Now routers can exchange the information about the multicast packets they have to receive among themselves.

Another important group of multicast protocols is the group of routing protocols. These protocols allow the routers to exchange information about multicast groups and thus to build routes for each group. Examples for multicast routing protocols are protocol-independent multicast (PIM) [154, 155], distance-vector multicast routing protocol (DVMRP) [156], and multicast open shortest path first (MOSPF) [157].

IP-layer multicast has not been widely adopted by most commercial ISPs, and thus large parts of the Internet are still incapable of IP multicast more than a decade after the protocols were developed. As a result, the Multicast backbone (MBONE) was developed [158]. It consists of "islands" of multicast enabled networks in the Internet, connected through different types of tunnels. This concept has some drawbacks like the manual tunnel setup and the need for constant IP addresses. This is not feasible for the average Internet user. However, with the increasing acceptance of the Internet Protocol Television (IPTV) some providers started to use multicast for the transport of live video streams, at least internally. It remains to be seen whether the support of IP multicast will be increased due to such new technologies.

6.6.1 Application Layer Multicast

Application layer multicast (ALM) is independent from the multicast support of the underlying network. The multicast forwarding functionality is implemented exclusively at end systems. Logically, the end systems form an overlay network, and the goal of application layer multicast is to construct and maintain an efficient overlay for data transmission. Since application layer multicast protocols cannot completely avoid the redundant transmission of data packets over the same link, they are less efficient than IP multicast. The advantages are that ALM systems do not require any modification of the underlying network components (e. g. routers) and can be implemented on the application layer without any special operating system support.

Figures 6.12 and 6.13 show the differences of the data flow in IP multicast and application layer multicast networks. The solid black lines identify the physical network connecting hosts and routers, while the dashed lines denote the data packet flow.

The IP routers in Fig. 6.12 forward the multicast packets from the sender to the receivers and duplicate the data if needed. The routers must therefore support the IP multicast protocol. In Fig. 6.13, the peer-to-peer (P2P) overlay connections are identified by a dotted and dashed line. In this environment no specialised routers are necessary. The packets are sent in unicast mode. The virtual data flow follows the overlay network structure, which does not necessarily correspond with the underly-

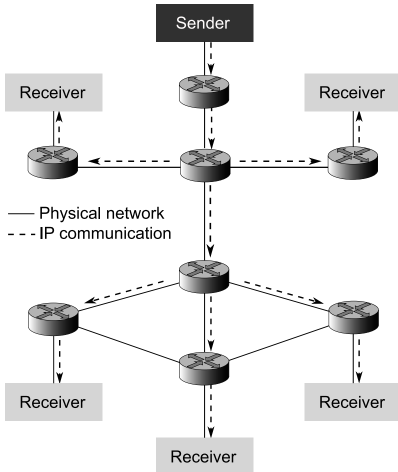

Fig. 6.12 IP multicast. The traffic is duplicated by the routers as needed.

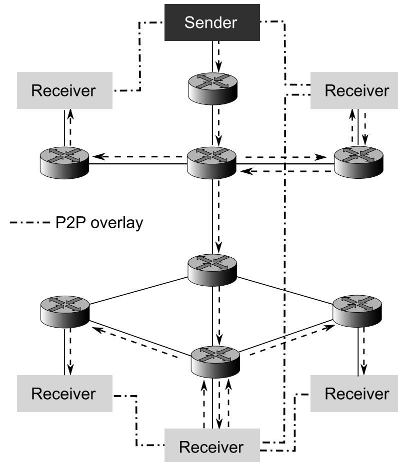

Fig. 6.13 Application layer multicast. Redundant data transmission cannot completely be avoided.

ing physical connections. However, the data is only replicated in the end systems, which are interconnected using unicast (P2P) links. Therefore, some packets are sent over the same link more than once. The efficiency of the ALM heavily depends on the overlay network construction and routing. With an optimal overlay topology, application layer multicast can approximate the efficiency of IP multicast.

6.6.2 Application Layer Multicast in the EuQoS System

Different ALM systems like Borg [159], VRing [160], Bayeux [161] or SplitStream [162] have been published over the past years. For the ALM support in the EuQoS system the combination Scribe/Pastry is used.

Pastry [163] is a scalable distributed object location and routing substrate for wide-area Peer-to-Peer applications. Nodes get an ID assigned when they join the Peer-to-Peer network. When a message needs to be sent to a certain ID, Pastry efficiently routes the message to the node with a node ID that is numerically closest to the ID of the message's destination. Pastry is self-organising, scalable and completely decentralised. It also takes node proximity (in terms of e2e delay) into account to minimise the distance messages are travelling.

Pastry uses a large ID space (2^{128} IDs), where hosts get random IDs assigned when joining the Peer-to-Peer network. The IDs are uniformly distributed over the whole ID space. This random assignment of IDs does not take locality nor Quality of Service requirements into account.

Pastry reliably routes messages identified by a key to the peer with the numerically closest ID to the key. Routing uses less than $\lceil log_{2^b} N \rceil$ steps on average, where N is the amount of nodes in the pastry network and b is typically a parameter with the value 4. Pastry guarantees eventual delivery unless $l/2$ or more nodes with an adjacent ID fail at the same time, with l, an even number parameter, being typically equal to 16. Pastry holds a routing table for each node with the size of $(2^b - 1)\lceil log_{2^b} N \rceil + l$ entries.

The routing tables are organised into $\lceil log_{2^b} N \rceil$ rows with each $2^b - 1$ entries. The entries of row n of a host's routing table point to other nodes, which share the same first n digits of their ID with the host itself, but the digit at position $n+1$ has one of the $2^b - 1$ possible values different from the digit at position $n+1$ of the host's ID. Each entry in the routing table consists of the node's ID and its corresponding IP address. Additionally, each node maintains a list of nodes (IDs and IP addresses) of the numerically closest hosts in its leaf set ($l/2$ entries for the larger and $l/2$ entries for the lower IDs). A message is routed to the closest (in terms the network latency) host found in the host's routing table whose ID matches the message's key prefix.

Figure 6.14 shows a simplified example of how Pastry routing works. A message with the key *e8cd* is routed from a peer with ID *3d1f* to the peer *e8ca*, which is numerically closest to the message key. On each hop from the source peer to the destination peer the message is sent to a peer whose ID matches more digits of the message key prefix as it did match at the hop before. For the first routing hop starting from peer *3d1f* the message is sent to peer *e2ce*, which shares the first digit *e* of the message key. At the second hop, the message is routed to the peer with ID *e831*, which shares the first two digits *e8*. Finally, it is sent to peer *e8ca* which is the peer closest to the message key and shares the first three digits *e8c* with the key.

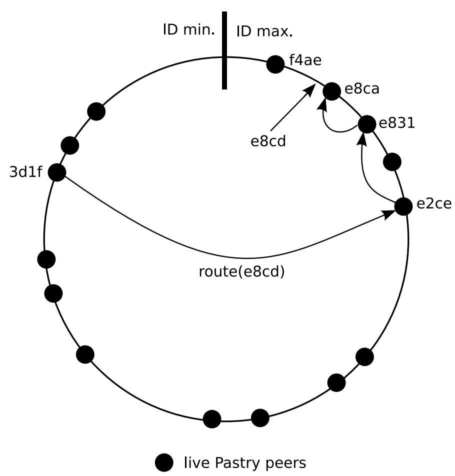

Fig. 6.14 Routing a message from peer *3d1f* to peer *e8cd*.

Scribe [164] builds on top of Pastry and is a large-scale decentralised Application Level Multicast infrastructure and supports a large number of groups and a potentially large number of members for each group. Scribe balances the load on nodes to achieve short delays and less link stress.

Any Scribe node can join any multicast group (or topic in Scribe terminology) at any time. For each topic, one node is designated to disseminate the topic data in the Pastry network. The node that is the root of the topic distribution tree has the ID numerically closest to the topic ID. Scribe offers best-effort delivery of the multicast data without guaranteeing that the order of the packets is maintained. The multicast or topic tree is built using a scheme similar to reverse-path-forwarding. A Scribe node, subscribing to a certain topic, sends a join message for this topic ID. This message is routed using Pastry's routing mechanism towards the topic's root. The next node, to which the join message is routed, remembers that the node sending the join message is interested in data for this topic. If this intermediate node, called a forwarder, has not already joined this topic, it will itself send a join message to the same topic. This process is repeated until a node is reached that has already joined the topic or is the root for the topic. Data dissemination within a topic is done from the root node of the topic towards the leaf nodes by following all reverse paths to the leaves. A side effect of this approach is that Scribe nodes forwarding messages for a certain topic have not necessarily interest in this topic.

6.6.3 Multicast Middleware

The EuQoS Multicast Middleware (MM in 6.15) [165, 166] is a solution to bridge application layer multicast and IP multicast. It provides a standard IP multicast interface for the applications on the sender and receiver side and uses Application Layer Multicast for transporting the data.

The Multicast Middleware can be used with any ALM network, which offers the standard multicast operations (subscription to a multicast group, receiving and sending multicast data). The typical P2P ALM network tries to approximate the efficiency of IP multicast communication regarding link stress by using unicast communication. As discussed earlier, ALM is not able to totally avoid sending redundant data over the same physical link as IP multicast can.

The overlay network is usually built in a topology aware manner. Therefore, peers that are "close" to each other in terms of communication latency are directly connected. The P2P links are constantly monitored, which allows reacting to failures in network communication or to failures of neighbour peers.

Eliminating the requirement for multicast support by the underlying network makes the use of Application Layer Multicast feasible for any kind of Internet users. The disadvantage of the ALM is the lack of standardisation. Each implementation has its own API and addressing scheme. This prohibits already existing multicast-aware applications from using the ALM.

The IP multicast interface for the applications is usually offered by the operating system. The operating system on the other side communicates with a multicast enabled router in the local network using IGMP as signalling protocol. Sending IP multicast traffic is not different from sending IP unicast traffic. The only difference is the reserved address range, which denotes different multicast groups (groups of multicast traffic receivers). On the link layer, multicast traffic is handled differently. For example, in Ethernet the IP packets with a multicast group as a destination address get an Ethernet multicast address assigned.

To provide an IP multicast interface for the whole system (including services integrated in the operating system's kernel), the Multicast Middleware uses a virtual Ethernet device (also known as TAP device—a software analogy of a wire tap). The TAP interface is a special kind of network interface, which is seen by the operating system as a normal Ethernet device. However, instead of forwarding the Ethernet frames to a hardware device, the TAP interface forwards the received Ethernet frames to a user-space process. On the other side, the TAP interface forwards all Ethernet frames received from the user-space process as incoming frames to the operating system's kernel. TAP support exists for all major operating systems such as UNIX/Linux, MacOS X and WIN32.

Using a TAP interface and the Multicast Middleware makes processing of multicast traffic transparent to all applications. This includes the multicast functionality integrated in the operating system's kernel. This approach does also not require any modification of application code. Any IP multicast application can be supported transparently. Multicast traffic originating from an end system can be routed through the TAP device. This device forwards the packets (encapsulated in Ethernet frames) to a user-space process (the Multicast Middleware) for processing. The Multicast Middleware acts as a multicast router by implementing IGMP and transporting the multicast data.

IP multicast enabled applications must subscribe to different multicast groups to receive video broadcast announcements and audio/video streams. The multicast group subscription is usually a system call, which instructs the operating system's kernel to send IGMP membership report messages to the IP multicast router. In our case, the IGMP membership reports are sent via the TAP interface to the Multicast Middleware. The Multicast Middleware interprets the IGMP membership reports and notifies the neighbour peers about the changes in the multicast routing table. This information (depending on the multicast routing protocol used in the overlay network) is propagated to other peers.

After a data packet has been sent by the application, it is forwarded by the operating system's kernel to the appropriate multicast enabled network device (in this case the TAP device). The Multicast Middleware process receives the outgoing multicast traffic via the TAP device. The received multicast traffic is then encapsulated into application layer multicast messages. The IP multicast destination address of the packets is translated into ALM addresses to which the messages are sent. Figure 6.15 shows the message flow for sending and receiving data with the Multicast Middleware. The application (APP) is running on both end-systems for sending/receiving the data stream. It uses the IP multicast interface of the Multicast Middleware that hides the ALM layer.

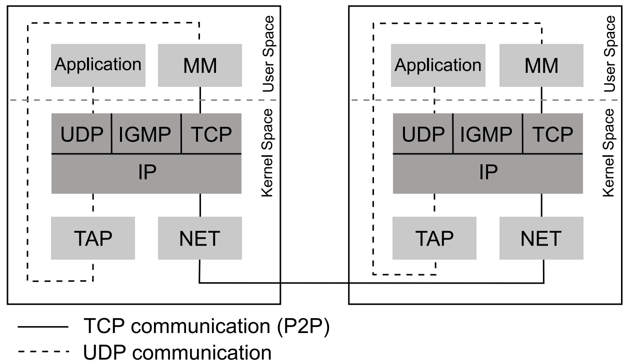

Fig. 6.15 Sending and receiving data using the Multicast Middleware (MM).

After receiving an encapsulated IP multicast packet by ALM, the Multicast Middleware encapsulates the IP multicast packet into an Ethernet frame. The Multicast Middleware then sends the Ethernet frame via the TAP interface to the operating system's kernel for processing. The operating system's kernel delivers the data to the application.

For the multicast data transport, any ALM protocol may be used. However, the mapping of the IP multicast address space to the application layer address scheme might differ from one protocol to another.

Every IP multicast packet has a destination address out of the IP multicast address range. Most application layer multicast protocols implement their own addressing scheme. Depending on the protocol's addressing scheme this address range can be smaller, equal or larger than the IP multicast address range. In case of a larger or equal address range, multicast addresses can be mapped one-to-one to the application layer multicast addresses. For example the IP multicast address range can be mapped to a consecutive address range of the same size in the application layer multicast protocol's addressing scheme. In the case where the address range of the application layer multicast is smaller than the IP multicast address range, the IP multicast addresses must be projected to the application layer multicast address range.

IP packets can be encapsulated in Application Layer Multicast messages. If the length of an Application Layer Multicast message is larger than the IP packet length, the standard IP packet fragmentation can be applied to the packet in order to transport the packet through the overlay network. On reception of fragmented IP packets, the Multicast Middleware should be able to reassemble them and to deliver them to the TAP interface. The time to live (TTL) field of the transported packets should be reduced for each P2P hop. Packets with TTL=0 should not be forwarded.

6.6.4 Introducing QoS to Multicast Middleware

To satisfy the QoS requirements, the Multicast Middleware uses the EuQoS system to setup network level QoS for the unicast links of the overlay network. Since the QoS requirements of the end systems within one IP multicast group can be heterogeneous, it is necessary that the multicast tree is built in such way that the QoS requirements and capabilities of end-systems are considered.

It is required that the QoS classes can be ordered and that they are independent of the path length. Such QoS classes can contain parameters such as bandwidth, jitter and maximum packet loss, but all the possible QoS classes must be comparable. Also note that in general there is no total order for a combination of such parameters and that the QoS parameter for maximum delay is not yet supported in the EuQoS system.

To provide QoS guarantees such as bandwidth or jitter in a multicast tree each e2e path from the root to a leaf node in the multicast tree must have a monotonically decreasing QoS requirement. Figure 6.16 shows an example of such a multicast tree. The path indicated as well as all other e2e paths of this multicast tree hold the following property: the QoS requirements (denoted by the thickness of the lines) are the same or decreasing when following the intermediate hops from the root node to a leaf node.

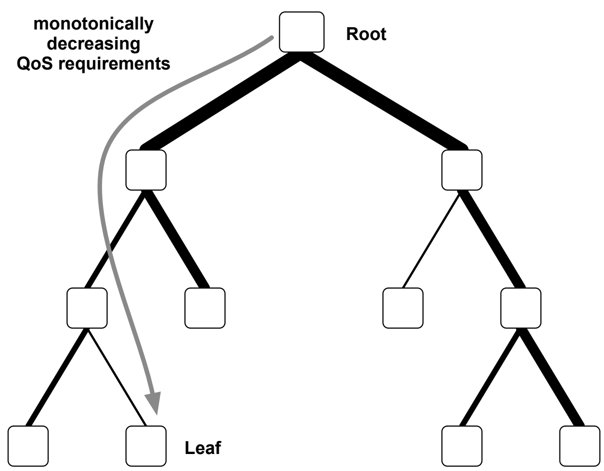

Fig. 6.16 Example of a multicast tree with monotonically decreasing QoS requirements from root to leaf nodes. Thickness of the lines represents the degree of the QoS requirement in terms of required bandwidth (thicker line = higher bandwidth requirement).

By analysing Scribe's multicast tree construction, it becomes clear that the constructed multicast tree does not necessarily hold this property. The reason for this is that the e2e path from a leaf to the root is more or less randomly chosen, due to random positioning of Pastry peers. Because Pastry's default ID assignment does not take QoS requirements of peers into account, the multicast trees constructed by Scribe are only by chance holding the described property. It is sufficient that only one link in an e2e path does not support this property to violate the QoS requirements for all nodes in the multicast tree below this link.

To enforce the construction of a QoS aware multicast tree using Scribe a dedicated Pastry P2P network is created for each multicast group. The reason for this is to have only peers interested in receiving the multicast data as potential forwarders. As a result, in this Pastry network only one topic exists. This topic ID is the highest possible topic ID. Since the QoS requirements of a peer can be higher than its QoS capabilities, the QoS class is chosen, which corresponds to the minimum of both.

As shown in Fig. 6.17, the ID space is partitioned into segments, one segment for each QoS class. Here, best-effort service is also considered to be a QoS class. The order of segments depends on the order of the QoS classes. The best-effort QoS class is located in the lowest segment and the highest QoS class is located in the highest segment. The assignment of IDs to joining peers depends on their QoS requirements/capabilities. The peer ID is randomly chosen within the corresponding segment of the ID space for the peer's QoS requirements/capabilities.

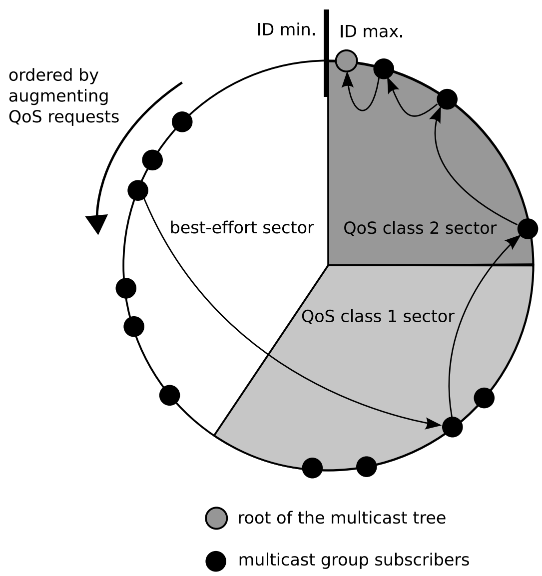

Fig. 6.17 QoS aware distribution of peer IDs for Pastry.

There are different possibilities on how large the segments should be. They do not necessarily have to be all of the same size and can for example decrease in size towards the root ID. The partitioning strategy has an impact on the construction of the multicast trees and therefore on how well and evenly balanced the overall traffic load will be distributed among the participating peers.

The routing path from a peer with a lower ID to a peer with a higher ID always contains peers with increasing IDs. Because the root node of the multicast tree has the highest possible Pastry ID, the routing should always use peers with increasing Pastry IDs for the hops on its path from leaf nodes towards the root node.

By assigning peer IDs proportional to the peer's QoS requirements, a construction of Scribe multicast trees, holding the decreasing QoS requirement property for each e2e path from the root to the leaves is ensured. For each node on the path from the root node to a leave node, the QoS requirement of the intermediate node is the same or lower than the one of its parent node.

6.7 Telemedicine Application

6.7.1 Telemedicine – the Case for Application-Driven QoS

The concept of utility, as a measure of the perceived value or benefit provided by QoS, is crucial to characterise the dependence of applications from QoS. For example, the utility of a standard mobile phone call is, up to a certain limit, relatively immune to QoS level variations – as long as end users are able to communicate in reasonable conditions. A certain level of QoS degradation is usually acceptable. Telemedicine applications are at the other end of the spectrum – near-perfect conditions are required by several application components (namely, real-time medical video), and tolerance to fluctuations of quality as a result of network load is minimal. The widespread deployment of such applications over public IP networks is often hindered by the limited capability of service providers to guarantee the strict fulfilment of those requirements.

Given this wide range of application characteristics and requirements, a challenge for service providers is how to handle such diversity both satisfactorily and efficiently. In theory, QoS could be handled by providing the most demanding QoS level to all customers and applications at all times - obviously, this would be economically unfeasible. Another solution could be the static allocation of specific QoS profiles to selected customers, thus ensuring that the required QoS treatment would be provided by the network to those customers in all circumstances. In either case, flexibility and cost effectiveness would be quite poor. The dynamic application-driven approach proposed by EuQoS described below provides a solution to this problem.

6.7.2 Overview of Medigraf

Medigraf is a real-time H.323-based telemedicine application including a video-conference, collaborative facilities and an embedded multimedia repository to store patient data, medical images and reports. A typical Medigraf screen-shot is shown in Fig. 6.18.

The application is used in several scenarios – performing remote cooperative diagnosis on a regular basis, providing remote specialised healthcare assistance, enabling collaboration between healthcare professionals in several scenarios, such as emergency situations and remote online training. The application offers significant gains in terms of efficiency and cost minimisation and is a valuable tool to provide specialised healthcare to populations in rural and sparsely populated areas, where the permanent availability of medical specialists is economically unfeasible.

Fig. 6.18 Typical Medigraf screenshot

The utility of the Medigraf application depends on the strict fulfilment of QoS parameters. In fact, proper medical diagnosis is not compatible with less than optimal QoS conditions. On the one hand, video quality is crucial to enable a correct medical diagnosis. On the other hand, e2e synchronisation of the application graphical elements requires stringent delay and jitter parameters.

Five basic traffic types are supported by Medigraf:

- Audio: used for audio communication. G.711 (PCMA, PCMU), G.728, G.722 and G.723 codecs are supported.
- Video: used for face-to-face communication and transfer of moving images acquired from specialized medical equipment (e.g. echocardiography). H.261, H.263 codecs are supported with CIF (352x288), QCIF (176x144) or SQCIF (128x96) resolution.

- Data: used for file transfer, typically medical images.
- Synchronisation: used for e2e synchronisation of graphical elements, e.g. pointers; requires strict compliance of e2e delay and jitter.
- Application control: only minimal values of loss and delay are tolerated.

The difficulty to guarantee appropriate QoS conditions has been one of the issues affecting the deployment of Medigraf. In some cases, static pre-reservation of network resources in specific time slots has been the solution to circumvent this problem. Unfortunately, this approach is not scalable and, in many circumstances, not realistic. Perhaps most important of all, it is not viable in unplanned or emergency situations, which coincidentally constitute one of the scenarios in which Medigraf would be most valuable. Clearly, a solution capable of providing e2e guarantees on a dynamic “on-demand” basis would be an important added value to the application.

6.7.3 Medigraf Adaptation to EuQoS

One of the innovative aspects of the EuQoS solution for e2e QoS is the awareness and active participation of the application in the QoS control process. We call an application EuQoS-aware if it is capable to explicitly request network resources and to actively participate in the QoS negotiation process by means of explicit signalling. This “EuQoS-awareness” requires the adaptation of the application or terminal to incorporate signalling capabilities, in order to inter-work with the EuQoS system. Two basic scenarios can be considered to integrate applications in the EuQoS system, as illustrated in Fig. 6.19:

1. Adapted legacy application: this refers to an application which was enhanced with a software add-on named APP (see Fig6.19) to interwork with QCM. The approach followed in the case of Medigraf falls into this category, as described below.
2. Proxy adapter: the strategy in this case is to leave the application untouched and use an external proxy adapter.

As explained in section 6.2.2, QCM plays a pivotal role in the EuQoS architecture, as it provides a common standard interface between applications and the EuQoS system. Through QCM, applications are able to manage QoS-enabled sessions and handle session events coming from the EuQoS system.

An EuQoS-aware version of Medigraf has been used to demonstrate and validate the basic EuQoS concept of application-driven e2e QoS. The EuQoS application-driven approach provides Medigraf with the capability to request and control the network resources it needs on a dynamic basis and provides a promising solution to enable the widespread use of medical applications over public IP network infrastructures.

The approach followed for adapting Medigraf is illustrated in Fig. 6.20. The original Medigraf application (“legacy” Medigraf) has been extended with a set of func-

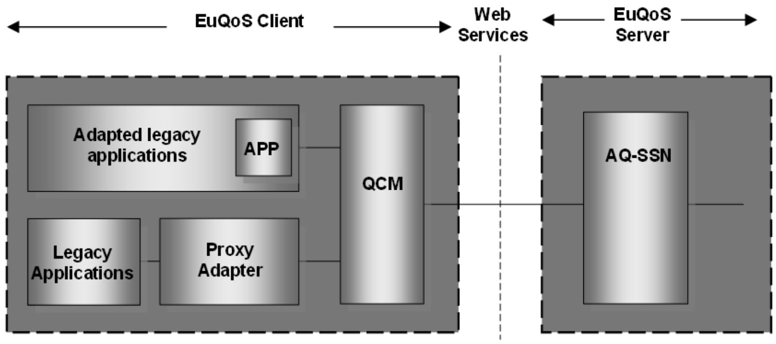

Fig. 6.19 Strategies for adaptation of applications

tions provided by the APP module, to enable inter-working with the EuQoS system through QCM.

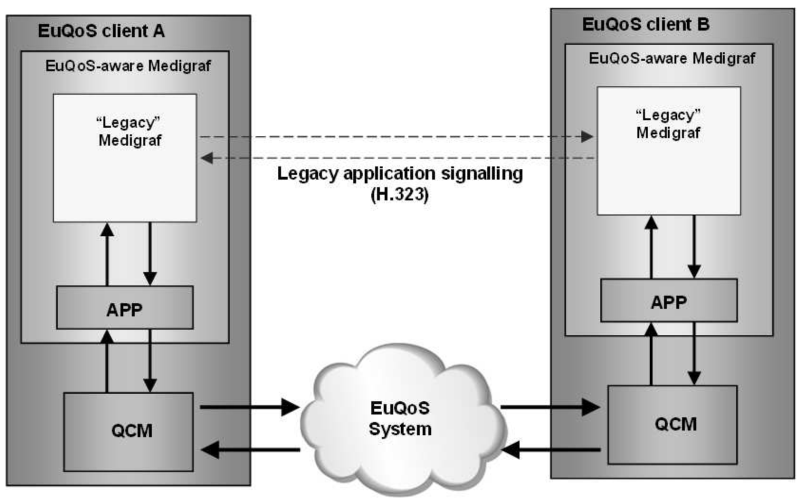

Fig. 6.20 Medigraf adaptation

APP incorporates the EuQoS awareness into the application and allows the communication with QCM and therefore the rest of the EuQoS system. This module has been integrated into the application to enable the invocation of the QCM methods ”performReservation” and ”closeReservation”.

Because Medigraf is natively based on H.323 signalling, a major design issue was how to deal with the coexistence of the legacy Medigraf H.323 signalling plane and the EuQoS signalling plane. To minimise the adaptation effort, it was decided to keep the H.323 plane untouched, moving the negotiation of QoS parameters to the EuQoS control plane and making sure that the result of the H.323 session setup (namely, codec characteristics, TCP/UDP port numbers) is consistent with the EuQoS control plane negotiation. The APP module must guarantee consistency and synchronisation between the two signalling planes, enforcing appropriate codec selection by means of hardware-specific functions, following the process illustrated in Fig. 6.21.

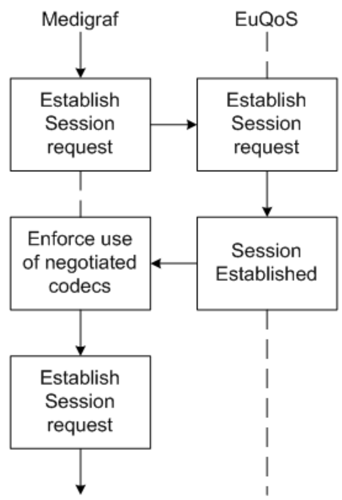

Fig. 6.21 Medigraf - EuQoS synchronisation

6.8 Conclusions

This chapter presented the global architecture developed in the EuQoS project for providing e2e QoS guarantees to Internet users. It addressed the problems of finding and providing e2e QoS paths between users connected through heterogeneous access network technologies.

A first prototype has been designed and implemented on a real testbed made up of GÉANT, the NRENs in each participating countries, and using different access network technologies, in particular WiFi, LAN, xDSL, UMTS, Satellite and MPLS.

The prototype implements all approaches and address all key problems, as application negotiation, application QoS on demand capability, QoS and technology independent signalling, admission control, network provisioning, resource management and layered integration of coherent protocols. All this is provided using network technology independent and dependent solutions.

The evaluation has shown that the global architecture is quite general, is able to integrate both a large set of technologies and a large set of independent ways of providing QoS. The network technology independent virtual layer proved to be quite efficient in terms of designing, handling and abstracting all real technologies and all dependent technology choices.

Efficiency and scalability have been proven, both in the access networks and in the core networks, specially in the admission control function, solved in the latter case by defining and using MPLS-based tunnels using PCE.

From the framework architecture adaptability and the results obtained, it follows that the EuQoS system is really generic and able to integrate a large set of solutions guaranteeing QoS in a unique e2e EuQoS architecture.

References

1. R. Steinmetz and K. Nahrstedt, *Multimedia Fundamentals: Medai Coding and Content Processing*, 2nd ed. Prentice-Hall PTR, 2002, vol. 1.
2. ——, *Multimedia Systems*, 1st ed. Berlin, Germany: Springer Verlag, 2004.
3. S. Floyd, "Connections with multiple congested gateways in packet-switched networks part 1: One-way traffic," *Computer Communications Review*, vol. 21, no. 5, pp. 30–47, October 1991.
4. M. Mathis, J. Semke, and J. Mahdavi, "The macroscopic behavior of the tcp congestion avoidance algorithm," *SIGCOMM Comput. Commun. Rev.*, vol. 27, no. 3, pp. 67–82, 1997.
5. ITU-T, "Itu t recommendation g.114: One-way transmission time." 2003.
6. ——, "Itu t recommendation g.1010: End-user multimedia qos categories." 2001.
7. T. Szigeti and C. Hattingh, *End-to-End QoS Network Design: Quality of Service in LANs, WANs, and VPNs (Networking Technology)*. CISCO Press, 2005.
8. S. Keshav, *An Engineering Approach to Computer Networking*. Addison Wesley, 1997.
9. R. Braden, D. Clark, and S. Shenker, "Integrated services in the internet architecture: an overview," RFC 1633, June 1994.
10. S. Shenker and J. Wroclawski, "General characterization parameters for integrated service network elements," RFC 2215, September 1997.
11. J. Wroclawski, "Specification of the Controlled-Load Network Element Service," RFC 2211 (Proposed Standard), September 1997. [Online]. Available: http://www.ietf.org/rfc/rfc2211.txt
12. S. Shenker, C. Partridge, and R. Guerin, "Specification of Guaranteed Quality of Service," RFC 2212 (Proposed Standard), September 1997. [Online]. Available: http://www.ietf.org/rfc/rfc2212.txt
13. J. Solomon, "Applicability Statement for IP Mobility Support," RFC 2005 (Proposed Standard), Oktober 1996. [Online]. Available: http://www.ietf.org/rfc/rfc2005.txt
14. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Service," RFC 2475 (Informational), Dezember 1998, updated by RFC 3260. [Online]. Available: http://www.ietf.org/rfc/rfc2475.txt
15. K. Nichols, S. Blake, F. Baker, and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," RFC 2474 (Proposed Standard), Dezember 1998, updated by RFCs 3168, 3260. [Online]. Available: http://www.ietf.org/rfc/rfc2474.txt
16. K. Nichols and B. Carpenter, "Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification," RFC 3086 (Informational), April 2001. [Online]. Available: http://www.ietf.org/rfc/rfc3086.txt
17. K. Nichols, V. Jacobson, and K. Poduri, "A per-domain behavior for circuit emulation in ip networks," *SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 2, pp. 71–83, 2004.
18. B. Davie, A. Charny, J. Bennet, K. Benson, J. L. Boudec, W. Courtney, S. Davari, V. Firoiu, and D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop Behavior)," RFC 3246 (Proposed Standard), März 2002. [Online]. Available: http://www.ietf.org/rfc/rfc3246.txt
19. J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski, "Assured Forwarding PHB Group," RFC 2597 (Proposed Standard), June 1999, updated by RFC 3260. [Online]. Available: http://www.ietf.org/rfc/rfc2597.txt
20. J. Babiarz, K. Chan, and F. Baker, "Configuration Guidelines for DiffServ Service Classes," RFC 4594 (Informational), August 2006. [Online]. Available: http://www.ietf.org/rfc/rfc4594.txt
21. L. Breslau, E. W. Knightly, S. Shenker, I. Stoica, and H. Zhang, "Endpoint admission control: architectural issues and performance," in *SIGCOMM '00: Proceedings of the conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*. New York, NY, USA: ACM, 2000, pp. 57–69.
22. K. Ramakrishnan and S. Floyd, "A Proposal to add Explicit Congestion Notification (ECN) to IP," RFC 2481 (Experimental), Januar 1999, obsoleted by RFC 3168. [Online]. Available: http://www.ietf.org/rfc/rfc2481.txt

23. C. Cetinkaya, V. Kanodia, and E. W. Knightly, "Scalable services via egress admission control," *IEEE Transactions on Multimedia*, vol. 3, no. 1, pp. 69–81, 2001. [Online]. Available: citeseer.ist.psu.edu/cetinkaya01scalable.html
24. S. McCanne, V. Jacobson, and M. Vetterli, "Receiver-driven layered multicast," in *ACM SIGCOMM*, vol. 26, no. 4. New York: ACM Press, Aug. 1996, pp. 117–130.
25. V. Paxson, "Strategies for Sound Internet Measurement," in *4rth Internet Measurements Conference*, 2004.
26. V. Paxson, G. Almes, J. Mahdavi, and M. Mathis, "Framework for IP Performance Metrics," RFC 2330, May 1998.
27. J. Mahdavi and V. Paxson, "IPPM Metrics for Measuring Connectivity," RFC 2678, Sept. 1999.
28. G. Almes, S. Kalidindi, and M. Zekauskas, "A One-way Delay Metric for IPPM," RFC 2679, Sept. 1999.
29. I.-T. R. Y.1541, "Network Performance Objectives fo IP-based Services," Review Jan 2005.
30. G. Almes, S. Kalidindi, and M. Zekauskas, "A Round-trip Delay Metric for IPPM," RFC 2681, Sept. 1999.
31. ——, "A One-way Packet Loss Metric for IPPM," RFC 2680, Sept. 1999.
32. C. Demichelis and P. Chimento, "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)," RFC 3393, Nov. 2002.
33. M. Mathis and M. Allman, "A Framework for Defining Empirical Bulk Transfer Capacity Metrics," RFC 3148, July 2001.
34. I.-T. R. P.800.1, "Mean Opinion Score (MOS) terminology," 2003.
35. A. Takahashi, H. Yoshino, and N. Kitawaki, "Perceptual qos assessment technologies for voip," *IEEE/Comm. Mag*, vol. 42, no. 7, pp. 28–34, 2004.
36. I.-T. R. G.107, "The E-model, a computational model for use in transmission planning," 2003.
37. I.-T. R. P.910, "Subjective video quality assessment methods for multimedia applications," 1999.
38. I.-T. R. P.862, "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," 2001.
39. D. L. Mills, "Network Time Protocol (Version 3) Specification, Implementation and Analysis," RFC 1305, Mar. 1992.
40. D. Veitch, S. Babu, and A. Pàsztor, "Robust synchronization of software clocks across the internet," in *Internet Measurement Conference (IMC '04)*, 2004.
41. N. Duffeld, "Sampling for passive internet measurement: A review," *Statistical Science*, vol. 19, no. 3, pp. 472–498, 2004.
42. J. Quittek, T. Zseby, B. Claise, and S. Zander, "Requirements for IP Flow Information Export (IPFIX)," RFC 3917, Oct. 2004.
43. M. Crovella and B. Krishnamurthy, *Internet Measurement. Infrastructure, Traffic, and Applications*. John Wiley & Sons, Ltd, 2006.
44. D. Harrington, R. Presuhn, and B. Wijnen, "An architecture for describing simple network management protocol (snmp) management frameworks," RFC 3411, Dec. 2002.
45. "Lobster - Large-scale Monitoring of Broadband Internet Infrastructures," http://www.ist-lobster.org.
46. N. G. Duffield and M. Grossglauser, "Trajectory sampling for direct traffic observation," *IEEE/ACM Trans. Netw.*, vol. 9, no. 3, pp. 280–292, 2001.
47. T. Zseby, S. Zander, and G. Carle, "Evaluation of Building Blocks for Passive One-Way-Delay Measurements," in *Passive and Active Measurements Conference*, 2001.
48. S. B. Moon, "Measurement and Analysis of End-to-end Delay and Loss in the Internet," Ph.D. dissertation, University of Massachusetts, 2000.
49. J. Strauss, D. Katabi, and F. Kaashoek, "A measurement study of available bandwidth estimation tools," in *Internet Measurement Conference (IMC)*, 2003.

50. L. Lao, C. Dovrolis, and M. Y. Sanadidi, "The probe gap model can underestimate the available bandwidth of multihop paths," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 5, pp. 29–34, 2006.
51. B. Melander, M. Bjorkman, and P. Gunningberg, "A New End-to-End Probing and Analysis Method for Estimating Bandwidth Bottlenecks," in *IEEE Globecom Global Internet Symposium*, 2000.
52. Q. Liu and J.-N. Hwang, "End-to-end available bandwidth estimation and time measurement adjustment for multimedia qos," in *ICME '03: Proceedings of the 2003 International Conference on Multimedia and Expo - Volume 3 (ICME '03)*. Washington, DC, USA: IEEE Computer Society, 2003, pp. 373–376.
53. A. Pásztor and D. Veitch, "Active Probing using Packet Quartets," in *Internet Measurement Workshop*, 2002.
54. R. Jain and S. A. Routhier, "Packet Trains-Measurements and a New Model for Computer Network Traffic," *IEEE Journal of Selected Areas in Communications*, vol. SAC-4, no. 6, pp. 986–995, 1986.
55. C. Dovrolis, P. Ramanathan, and D. Moore, "Packet-dispersion techniques and a capacity-estimation methodology," *IEEE/ACM Trans. Netw.*, vol. 12, no. 6, pp. 963–977, 2004.
56. "PlanetLab: An open platform for developing, deploying, and accessing planetary-scale services," http://www.planet-lab.org/.
57. "Cooperative Association for Internet Data Analysis (CAIDA)," http://www.caida.org/.
58. "Active Measurement Project (AMP)," http://amp.nlanr.net/.
59. F. Georgatos, F. Gruber, D. Karrenberg, M. Santcroos, A. sanj, H. Uijterwaal, and R. Wilhem, "Providing Active Measurements as a Regular Service for ISP's," in *Proceedings of Passive and Active Measurement*, 2001.
60. "Evergrow Traffic Observatory Measurement InfrastruCture," http://www.etomic.org/.
61. "DIMES (Distributed Internet MEasurements & Simulations)," http://www.netdimes.org/.
62. "Ever-growing global scale-free networks, their provisioning, repair and unique functions," http://www.evergrow.org/.
63. S. Shalunov, B. Teitelbaum, A. Karp, J. W. Boote, and M. J. Zekauskas, "One-Way Active Measurements Protocol," RFC 4656, Sept. 2006.
64. S. McCanne and V. Jacobson, "The BSD Packet Filter: A New Architecture for User-level Packet Capture," in *USENIX Winter*, 1993.
65. "perfSONAR - PERFormance Service-Oriented Network monitoring ARchitecture," http://www.perfsonar.net.
66. I. Miloucheva, P. Gutierrez, D. Hetzer, A. Nassri, and M. Beoni, "Intermon architecture for complex QoS analysis in inter-domain environment based on discovery of topology and traffic impact," in *Inter-domain Performance and Simulation Workshop, Budapest*, March 2004.
67. "Passive Measurement and Analysis (PMA)," http://pma.nlanr.net/.
68. D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell, and J. McManus, "Requirements for Traffic Engineering Over MPLS," RFC 2702 (Informational), Sept. 1999. [Online]. Available: http://www.ietf.org/rfc/rfc2702.txt
69. E. Rosen, D. Tappan, G. Fedorkow, Y. Rekhter, D. Farinacci, T. Li, and A. Conta, "MPLS Label Stack Encoding," RFC 3032 (Proposed Standard), Jan. 2001, updated by RFCs 3443, 4182. [Online]. Available: http://www.ietf.org/rfc/rfc3032.txt
70. L. Andersson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas, "LDP Specification," RFC 3036 (Proposed Standard), Jan. 2001. [Online]. Available: http://www.ietf.org/rfc/rfc3036.txt
71. H. Smit and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)," RFC 3784 (Informational), June 2004, updated by RFC 4205. [Online]. Available: http://www.ietf.org/rfc/rfc3784.txt
72. D. Katz, K. Kompella, and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2," RFC 3630 (Proposed Standard), Sept. 2003, updated by RFC 4203. [Online]. Available: http://www.ietf.org/rfc/rfc3630.txt

73. R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, “Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification,” RFC 2205 (Proposed Standard), Sept. 1997, updated by RFCs 2750, 3936, 4495. [Online]. Available: http://www.ietf.org/rfc/rfc2205.txt
74. D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow, “RSVP-TE: Extensions to RSVP for LSP Tunnels,” RFC 3209 (Proposed Standard), Dec. 2001, updated by RFCs 3936, 4420, 4874. [Online]. Available: http://www.ietf.org/rfc/rfc3209.txt
75. B. Jamoussi, L. Andersson, R. Callon, R. Dantu, L. Wu, P. Doolan, T. Worster, N. Feldman, A. Fredette, M. Girish, E. Gray, J. Heinanen, T. Kilty, and A. Malis, “Constraint-Based LSP Setup using LDP,” RFC 3212 (Proposed Standard), Jan. 2002, updated by RFC 3468. [Online]. Available: http://www.ietf.org/rfc/rfc3212.txt
76. D. Grossman, “New Terminology and Clarifications for Diffserv,” RFC 3260 (Informational), Apr. 2002. [Online]. Available: http://www.ietf.org/rfc/rfc3260.txt
77. F. Le Faucheur, L. Wu, B. Davie, S. Davari, P. Vaananen, R. Krishnan, P. Cheval, and J. Heinanen, “Multi-Protocol Label Switching (MPLS) Support of Differentiated Services,” RFC 3270 (Proposed Standard), May 2002. [Online]. Available: http://www.ietf.org/rfc/rfc3270.txt
78. F. L. Faucheur and W. Lai, “Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering,” RFC 3564 (Informational), July 2003. [Online]. Available: http://www.ietf.org/rfc/rfc3564.txt
79. F. Le Faucheur and W. Lai, “Maximum Allocation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering,” RFC 4125 (Experimental), June 2005. [Online]. Available: http://www.ietf.org/rfc/rfc4125.txt
80. F. Le Faucheur, “Russian Dolls Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering,” RFC 4127 (Experimental), June 2005. [Online]. Available: http://www.ietf.org/rfc/rfc4127.txt
81. W. Lai, “Bandwidth Constraints Models for Differentiated Services (Diffserv)-aware MPLS Traffic Engineering: Performance Evaluation,” RFC 4128 (Informational), June 2005. [Online]. Available: http://www.ietf.org/rfc/rfc4128.txt
82. F. Le Faucheur, “Protocol Extensions for Support of Diffserv-aware MPLS Traffic Engineering,” RFC 4124 (Proposed Standard), June 2005. [Online]. Available: http://www.ietf.org/rfc/rfc4124.txt
83. R. Stewart, Q. Xie, K. Morneault, and H. Schwarzbauer, “Stream Control Transmission Protocol,” IETF, Tech. Rep. 2960, 2000.
84. T. Dierks and C. Allen, “The TLS Protocol,” IETF, Tech. Rep. 2246, 1999.
85. D. Katz, “Router Alert Option,” IETF, Tech. Rep. 2113, 1997.
86. M.Stiemerling, H.Tschofenig, C. Aoun, and E. Davies, “Nat/firewall nsis signaling layer protocol (nslp),” IETF,” internet draft, 2006.
87. D.Durham, J.Boyle, R.Cohen, S.Herzog, R.Rajan, and A.Sastry, “The COPS (Common Open Policy Service) Protocol,” IETF, Tech. Rep. 2748, 2000.
88. isox214, “ITU-T recommendation X.214 (11/95) information technology,” Nov. 1995.
89. J. Postel, “Transmission control protocol: DARPA internet program protocol specification,” IETF, Request For Comments 793, 1981.
90. ——, “User datagram protocol (UDP),” IETF, Request For Comments 768, Aug. 1980.
91. H.Schulzrinne, S.Casner, R. Frederick, and V. Jacobson, “An extension to the Selective Acknowledgment (SACK) Options for TCP,” IETF, Task Force 3550, 2003.
92. V. Jacobson and R. Braden, “TCP extensions for long-delay paths,” IETF, Request For Comments 1072, Oct. 1988.
93. V. Jacobson, R. Braden, and D. Borman, “Tcp extensions for high performance,” IETF, Request For Comments 1323, 1992.
94. L. S. Brakno and L. L. Peterson, “TCP Vegas: End to End congestion Avoidance on a Global Internet,” *IEEE Journal on Selected Areas in Communications*, no. 13, pp. 1465–1480, 1995.
95. S. Floyd, T. Henderson, and A. Gurtov, “The newreno modification to tcp’s fast recovery algorithm,” IETF, Request For Comments 3782, Apr. 2004.
96. S. Floyd, J. Mahdavi, M. Mathis, and M. Podolsky, “An extension to the selective acknowledgement (SACK) option for TCP,” IETF, Request For Comments 2883, July 2000.

97. O. Ait-Hellal and E. Altman, “Analysis of tcp-vegas and tcp-reno,” in *Proc. of the IEEE International Conference on Communications - ICC*, 1997, pp. 495–499.
98. V. Jacobson, L. Peterson, L. Brakmo, and S. Floyd, “Problems with arizona’s vegas,” mailing list, end2end-tf, Tech. Rep., 1994, ftp://ftp.ee.lbl.gov/email/vanj.94mar14.txt.
99. L. Xu, K. Harfoush, and I. Rhee, “Binary increase congestion control for fast long-distance networks,” in *Proc. of IEEE INFOCOM*, Mar. 2004, pp. 2514–2524.
100. S. Floyd, “Highspeed tcp for large congestion windows,” IETF, Request For Comments 3649, Dec. 2003.
101. D. Katabi, M. Handley, and C. Rohrs, “Congestion control for high bandwidth-delay product networks,” in *Proc. of ACM SIGCOMM*, 2002.
102. Y. Xia, L. Subramanian, I. Stoica, and S. Kalyanaraman, “One more bit is enough,” in *Proc. of ACM SIGCOMM*, 2005.
103. K. Ramakrishnan, S. Floyd, and D. Black, “The addition of explicit congestion notification (ECN) to ip,” IETF, Request For Comments 3168, Sept. 2001.
104. P. Sinha, T. Nandagopal, N. Venkitaraman, R. Sivakumar, and V. Bharghavan, “WTCP: A reliable transport protocol for wireless wide-area networks,” *Wireless Networks*, vol. 8, no. 2-3, pp. 301–316, 2002.
105. M. Gerla, M. Y. Sanadidi, R. Wang, A. Zanella, C. Casetti, and S. Mascolo, “Tcp westwood: congestion window control using bandwidth estimation,” in *Proc. of IEEE GLOBECOM*, 2001, pp. 1698–1702.
106. C. P. Fu and S. C. Liew, “Tcp veno: Tcp enhancement for transmission over wireless access networks,” *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 2, pp. 216–229, Feb. 2003.
107. R. S. et Al., “Sockets API extensions for stream control transmission protocol (SCTP),” IETF, Internet Draft draft-ietf-tsvwg-sctpsocket-06.txt, Mar. 2003.
108. E. Kohler, M. Handley, and S. Floyd, “Datagram congestion control protocol (DCCP),” IETF, Request For Comments 4340, Mar. 2006.
109. S. Floyd and E. Kohler, “Profile for DCCP Congestion Control ID 2: TCP-like Congestion Control,” IETF, Request For Comments 4341, Mar. 2006.
110. S. Floyd, E. Kohler, and J. Padhye, “Profile for DCCP congestion control ID 3: TRFC congestion control,” IETF, Request For Comments 4342, Mar. 2006.
111. M. Handley, S. Floyd, J. Pahdye, and J. Widmer, “Tcp-friendly rate control (TFRC): Protocol specification,” IETF, Request For Comments 3448, Jan. 2003.
112. S. Iren, P. D. Amer, and P. T. Conrad, “The transport layer: tutorial and survey,” *ACM Computer Survey*, vol. 31, no. 4, pp. 360–404, 1999.
113. S. Floyd, “Congestion control principles,” IETF, Request For Comments 2914, Sept. 2000.
114. Y.-Q. Z. D. Wu, T. Hou, “Transporting real-time video over the internet: Challenges and approaches,” in *Proc. of the IEEE*, vol. 88, no. 12, Dec. 2000, pp. 1855–1875.
115. V. Jacobson, “Congestion avoidance and control,” in *Proc. of ACM SIGCOMM*, Stanford, CA, Aug. 1988, pp. 314–329.
116. R. Talluri, “Error-resilience video coding in the ISO MPEG-4 standard,” *IEEE Communications Magazine*, pp. 112–119, June 1998.
117. Y. Wang and Q.-F. Zhu, “Error control and concealment for video communication: A review,” in *Proc. of the IEEE*, vol. 86, no. 5, May 1998, pp. 974–997.
118. M. Mathis, J. Mahdavi, S. Floyd, and A. Romanow, “TCP selective acknowledgment options,” IETF, Request For Comments 2018, Oct. 1996.
119. J. Widmer, “Equation-based congestion control,” Diploma Thesis, University of Mannheim, Germany, Feb. 2000.
120. S. Floyd and K. Fall, “Promoting the use of end-to-end congestion control in the internet,” *IEEE/ACM Transactions on Networking*, vol. 7, no. 4, pp. 458–472, 1999.
121. P. Amer, C. Chassot, C. Connolly, P. Conrad, and M. Diaz, “Partial order transport service for MM and other application,” *IEEE/ACM Transactions on Networking*, vol. 2, no. 5, 1994.
122. L. Rojas-Cardenas, E. Chaput, L. Dairaine, P. Snac, and M. Diaz, “Video transport over partial order connections,” *Computer Networks*, vol. 31, no. 7, pp. 709–725, Apr. 1999.

123. J.-P. V. A. Farrel and J. Ash, “A Path Computation Element (PCE)-Based Architecture,” The Internet Society, Request For Comments 4655, Aug. 2006.
124. J. Enríquez, M. A. Callejo, and et al., “EuQoS Architecture Deliverable D122,” 2007.
125. A. Beben, “EQ-BGP: an efficient inter-domain QoS routing protocol,” in *AINA ’06: Proceedings of the 20th IEEE International Conference on Advanced Information Networking and Applications*. Los Alamitos, CA, United States: IEEE Computer Society, Apr. 2006, pp. 560–564.
126. X. Masip-Bruin, M. Yannuzzi, R. Serral-Gracia, J. Domingo-Pascual, J. Enriquez-Gabeiras, M. Callejo, M. Diaz, F.Racaru, G. Stea, E. Mingozzi, A. Beben, W. Burakowski, E. Monteiro, and L. Cordeiro, “The EuQoS System: A solution for QoS Routing in Heterogeneous Networks,” *IEEE Communications Magazine*, vol. 45, no. 2, Feb. 2007.
127. Y. Rekhter, T. Li, and S. Hares, “A Border Gateway Protocol 4 (BGP-4),” The Internet Society, Request For Comments 4271, Jan. 2006.
128. L. Baresse and et al., “Integrated EuQoS system Software architecture for application use cases and API for application driving Deliverable D421.”
129. P. Krawiec and W. Burakowski, “Resource allocation strategies for new connections in qos multi-domain networks with signaling capabilities,” in *Proceedings of Australian Telecommunication Networks and Applications Conference 2007 (ATNAC 2007)*, Christchurch, New Zealand, Dec. 2007, pp. 337–342.
130. A. Beben and et al., “Definition of measurement and monitoring system for Phase 2 Deliverable D222.”
131. A. Bak, W. Burakowski, F. Ricciato, S. Salsano, and H. Tarasiuk, “A framework for providing differentiated QoS guarantees in IP-based network,” *Computer Communications*, vol. 26, no. 4, pp. 327–337, 2003.
132. C. Brandauer, W. Burakowski, M. Dabrowski, B. Koch, and H. Tarasiu, “AC algorithms in AQUILA QoS IP network,” *European Transactions on Telecommunications*, vol. 16, no. 3, pp. 225–232, 2005.
133. M. Dabrowski, G. Eichler, M. Fudala, D. Katzengruber, T. Kilkanen, N. Miettinen, H. Tarasiuk, and M. Titze, “Evaluation of the AQUILA Architecture: Trial Results for Signalling Performance, Network Services and User Acceptance,” in *Art-QoS*, 2003, pp. 218–233.
134. K.Chan, J.Babiarz, and F.Baker, “Aggregation of DiffServ Service Classes,” Transport Area Working Group,” Internet-Draft, Nov. 2007.
135. “Nexuiz project,” www.alientrap.org/nexuiz/.
136. J. M. Batalla and R. Janowski, “Provisioning dedicated class of service for reliable transfer of signaling traffic,” in *International Teletraffic Congress*, ser. Lecture Notes in Computer Science, L. Mason, T. Drwiega, and J. Yan, Eds., vol. 4516. Springer, 2007, pp. 853–864.
137. “Medigraf,” http://www.medigraf.pt/.
138. J. W. Roberts, U. Mocci, and J. T. Virtamo, Eds., *Broadband Network Teletraffic - Performance Evaluation and Design of Broadband Multiservice Networks: Final Report of Action COST 242*, ser. Lecture Notes in Computer Science. Springer, 1996, vol. 1155.
139. L. Kleinrock, *Queueing Systems, Volume 2, Computer Applications*. John Wiley & Sons Inc, 1976.
140. H. Tarasiuk, R. Janowski, and W. Burakowski, “Admissible traffic load of real time class of service for inter-domain peers,” in *Proceedings of Joint International Conference on Autonomic and Autonomous Systems and International Conference on Networking and Services (ICAS/ICNS 2005)*. Papeete, Tahiti, French Polynesia: IEEE Computer Society, Oct. 2005.
141. ——, “Application of Admission Control and Traffic Shaping for providing TCP Throughput Guarantees,” in *Proceedings of International Workshop To-QoS’2006 in conjunction with IFIP 2006 Networking Conference (ed. W.Burakowski)*, Coimbra, Portugal, May 2006, pp. 163–172.
142. “IEEE standards for local and metropolitan area networks. Virtual bridged local area networks,” *IEEE Std 802.1Q, 2003 Edition (Incorporates IEEE Std 802.1Q-1998, IEEE Std 802.1u-2001, IEEE Std 802.1v-2001, and IEEE Std 802.1s-2002)*, 2003.
143. “IEEE Standard for Local and metropolitan area networks Media Access Control (MAC) Bridges,” *IEEE Std 802.1D-2004 (Revision of IEEE Std 802.1D-1998)*, 2004.

144. M. Carmo, J. S. Silva, E. Monteiro, P. Simões, and F. Boavida, "Ethernet QoS Modeling in Emerging Scenarios," in *Proceedings of 3rd International Workshop on Internet Performance, Simulation, Monitoring and Measurement (IPS-MoMe 2005)*. Warsaw, Poland: IST MoMe Cluster, Mar. 2005, pp. 90–96.
145. M. Carmo, B. Carvalho, J. S. Silva, E. Monteiro, P. Simões, M. Curado, and F. Boavida, "NSIS-based Quality of Service and Resource Allocation in Ethernet Networks," in *Proceedings of 4th International Conference on Wired/Wireless Internet Communications 2006*, 2006.
146. M. Carmo, J. S. Silva, and E. Monteiro, "EuQoS approach for Resource Allocation in Ethernet Networks," *International Journal of Network Management*, vol. 17, no. 5, pp. 373–388, sep / oct 2007.
147. "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements," *IEEE Std 802.11e-2005*, 2005.
148. E. Lochin and L. D. A. Jourjon, "gtfrc, a tcp friendly qos-aware rate control for diffserv assured service," *Springer Telecommunication Systems*, vol. 10.1007/s11235-006-9004-2, ISSN: 1018-4864 (Print) 1572-9451 (Online), September 2006 2006.
149. E. Exposito, P. Sénac, and M. Diaz, "Compositional architecture pattern for qos-oriented communication mechanisms," in *MMM*, Y.-P. P. Chen, Ed. IEEE Computer Society, 2005, pp. 413–420.
150. B. Quinn and K. Almeroth, "Ip multicast applications: Challenges and solutions," The Internet Society, Request For Comments 3170, Sept. 2001.
151. Z. Albanna, K. Almeroth, D. Meyer, and M. Schipper, "Iana guidelines for ipv4 multicast address assignments," The Internet Society, Request For Comments 3171, Aug. 2001.
152. R. Hinden and S. Deering, "Internet protocol version 6 (ipv6) addressing architecture," The Internet Society, Request For Comments 3513, Apr. 2003.
153. B. Cain, S. Deering, I. Kouvelas, B. Fenner, and A. Thyagarajan, "Internet group management protocol, version 3," The Internet Society, Request For Comments 3376, Oct. 2002.
154. A. Adams, J. Nicholas, and W. Siadak, "Protocol independent multicast - dense mode (pim-dm): Protocol specification (revised)," The Internet Society, Request For Comments 3973, Jan. 2005.
155. B. Fenner, M. Handley, H. Holbrook, and I. Kouvelas, "Protocol independent multicast - sparse mode (pim-sm): Protocol specification (revised)," The Internet Society, Request For Comments 4601, Aug. 2006.
156. D. Waitzman, C. Partridge, and S. Deering, "Distance vector multicast routing protocol," Network Working Group, Request For Comments 1075, Nov. 1988.
157. J. Moy, "Multicast extensions to ospf," Network Working Group, Request For Comments 1584, Mar. 1994.
158. K. Savetz, N. Randall, and Y. Lepage, *MBONE: Multicasting Tomorrow's Internet*. John Wiley & Sons Inc, 1996.
159. R. Zhang and Y. C. Hu, "Borg: A hybrid protocol for scalable application-level multicast in peer-to-peer networks," in *NOSSDAV '03: Proceedings of the 13th international workshop on Network and operating systems support for digital audio and video*, ACM. New York, NY, USA: ACM Press, June 2003, pp. 172–179.
160. A. Sobeih, W. Yurcik, and J. C. Hou, "Vring: A case for building application-layer multicast rings (rather than trees)," in *MASCOTS '04: Proceedings of the The IEEE Computer Society?s 12th Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems (MASCOTS'04)*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 437–446.
161. S. Q. Zhuang, B. Y. Zhao, A. D. Joseph, R. H. Katz, and J. D. Kubiatowicz, "Bayeux: An architecture for scalable and fault-tolerant wide-area data dissemination," in *NOSSDAV '01: Proceedings of the 11th international workshop on Network and operating systems support for digital audio and video*. Port Jefferson, New York, United States: ACM Press, 2001, pp. 11–20.

162. M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, “Splitstream: High-bandwidth multicast in a cooperative environment,” in *19th ACM Symposium on Operating Systems Principles (SOSP’03)*, Bolton Landing, New York, USA, oct 2003.
163. A. Rowstron and P. Druschel, “Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems,” in *IFIP/ACM International Conference on Distributed Systems Platforms (Middleware)*, Heidelberg, Germany, nov 2001, pp. 329–350.
164. M. Castro, P. Druschel, A.-M. Kermarrec, and A. Rowstron, “Scribe: A large-scale and decentralized application-level multicast infrastructure,” *IEEE Journal on Selected Areas in Communication (JSAC)*, vol. 20, no. 8, oct 2002.
165. M. Brogle, D. Milic, and T. Braun, “Qos enabled multicast for structured p2p networks,” in *4th IEEE Consumer Communications and Networking Conference*, Las Vegas, NV, USA, Jan. 2007.
166. D. Milic, M. Brogle, and T. Braun, “Video broadcasting using overlay multicast,” in *Seventh IEEE International Symposium on Multimedia (ISM 2005)*, Irvine, CA, USA, Dec. 2005, pp. 515–522.
167. A. Sulistio, C. S. Yeo, and R. Buyya, “A taxonomy of computer-based simulations and its mapping to parallel and distributed systems simulation tools,” *Softw. Pract. Exper.*, vol. 34, no. 7, pp. 653–673, 2004.
168. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” *The Journal of Chemical Physics*, vol. 21, pp. 1087–1092, 1953.
169. T. Yung, J. Martin, M. Takai, and R. Bagrodia, “Integration of fluidbased analytical model with packet-level simulation for analysis of computer networks,” 2001.
170. C. Kiddle, R. Simmonds, C. Williamson, and B. Unger, “Hybrid packet/fluid flow network simulation,” in *PADS ’03: Proceedings of the seventeenth workshop on Parallel and distributed simulation*. Washington, DC, USA: IEEE Computer Society, 2003, p. 143.
171. J. Incera, R. Marie, D. Ros, and G. Rubino, “Fluidsim: a tool to simulate fluid models of high-speed networks,” *Perform. Eval.*, vol. 44, no. 1-4, pp. 25–49, 2001.
172. B. Liu, D. R. Figueiredo, Y. Guo, J. F. Kurose, and D. F. Towsley, “A study of networks simulation efficiency: Fluid simulation vs. packet-level simulation,” in *INFOCOM*, 2001, pp. 1244–1253.
173. “Glomosim,” http://pcl.cs.ucla.edu/projects/glomosim/.
174. “Qualnet,” http://www.scalable-networks.com/.
175. “Swans,” http://jist.ece.cornell.edu/.
176. “Ssfnet,” http://www.ssfnet.org/homePage.html.
177. “Omnet++,” http://www.omnetpp.org/.
178. “ns-2,” http://www.isi.edu/nsnam/ns/index.html.
179. “Vint - virtual internetwork testbed project,” http://www.isi.edu/nsnam/vint/index.html.
180. K. Fall and K. Varadhan, “The ns manual (formerly ns notes and documentation),” 2002.
181. F. J. Ros and P. M. Ruiz, “Implementing a new manet unicast routing protocol in ns2,” Dept. of Information and Communications Engineering University of Murcia, Tech. Rep., 2004.
182. “Vmware software (virtual machine),” *http://www.vmware.com*, 2005.
183. V. J. Easton and J. H. McColl, “Statistics glossary v1.1,” *http://www.cas.lancs.ac.uk/glossaryv1.1/main.html*, vol. 2005, 2005.
184. M. Allman, A. Caldwell, and S. Ostermann, “One: The ohio network emulator,” *TR-19972*, 1997.
185. L. Rizzo, “Dummynet: a simple approach to the evaluation of protocols,” *ACM Computer Communication Review*, 1997.
186. M. Carson and D. Santay, “Nist net: A linux-based network emulation tool,” *ACM Computer Communication Review*, 2003.
187. S. Hemminger, “Netem, network emulator,” *http://developer.osdl.org/shemminger/netem/*, 2005.
188. P. Zheng and L. M. Nil, “Empower: A network emulator for wireline and wireless networks,” in *IEEE Infocom*, San Francisco, California, USA, 2003.

189. R. Buyya, D. Abramson, and J. Giddy, “A case for economy grid architecture for service oriented grid computing,” in *10th IEEE International Heterogeneous Computing Workshop*, San Francisco, California, USA, 2001.
190. B. White, J. Lepreau, L. Stoller, R. Ricci, S. Guruprasad, M. Newbold, M. Hibler, C. Barb, and A. Joglekar, “An integrated experimental environment for distributed systems and networks,” in *5th symposium on Operating systems design and implementation*, Boston, USA, 2002.
191. CNRS, “Grid 5000,” *http://www.grid5000.org*, 2003.
192. L. Peterson, T. Anderson, D. Culler, and T. Roscoe, “A blueprint for introducing disruptive technology into the internet,” in *First Workshop on Hot Topics in Networking (HotNets-I)*, Princeton, New Jersey, USA, 2002.
193. D. Herrscher and K. Rothermel, “A dynamic network scenario emulation tool,” in *11th International Conference on computer communications and networks*, Florida, USA, 2002.
194. “Network node emulation and method of node emulation,” *European patent*, 2002.
195. S. Dawson and F. Jahanian, “Probing and fault injection of distributed protocols implementations,” in *International Conference on Distributed Computer Systems*, 1995.
196. M. Zec and M. Mikuc, “Operating system support for integrated network emulation in imunes,” in *First Workshop on Operating System and Architectural Support for the on demand IT InfraStructure*, Boston, USA, 2004.
197. X. W. Huang, R. Sharma, and S. Keshav, “The entrapid protocol development environment,” in *IEEE Infocom*, Boston, USA, 1999.
198. E. Savage and D. Wethereall, “Alpine, a user level infrastructure for network development environment,” in *IEEE Infocom*, Boston, USA, 1999.
199. F. Baumgertner, T. Braun, and B. Bhargava, “Virtual routers: a tool for emulating ip routers,” in *27th Annual IEEE Conference on Local Computer Networks*, 2002.
200. B. Noble, M. Satyanarayanan, D. Narayanan, J. E. Tilton, J. Flinn, and K. R. Walker, “Trace-based mobile network emulation,” in *ACM SIGCOMM*, Cannes, France, 1997.
201. NS-Group, “The network simulator - ns-2,” *http://www.isi.edu/nsnam/ns*, 1989.
202. K. Fall, “Network emulation in the vint/ns simulator,” in *Fourth IEEE Symposium on Computers and Communications*, 1999.
203. C. Kiddle, R. Simmonds, and B. Unger, “Improving scalability of network emulation through parallelism and abstraction,” in *38th Annual Simulation Symposium (ANSS’05)*, 2005.
204. M. Gineste, H. Thalmensy, L. Dairaine, P. Senac, and M. Diaz, “Active emulation of a dvb-rcs satellite link in an end-to-end qos-oriented heterogeneous network,” in *23rd AIAA International Communication Satellite Systems (ICSSC)*, Rome, Italy, 2005.
205. ETSI, “Dvb-rcs: Digital video broadcasting (dvb); interaction channel for satellite distribution systems,” vol. EN 301 790, 2003.
206. ——, “Digital video broadcasting (dvb); dvb specification for data broadcasting,” vol. EN 301 192, 1999.
207. ISO/IEC, “Information technology – generic coding of moving pictures and associated audio: Systems,” vol. 13818-1, 1994.
208. ——, “Information technology - generic coding of moving pictures and associated audio information - part 6: Extensions for dsm-cc,” vol. 13818-6, 1998.
209. S.Kota and M. Marchese, “Quality of service for satellite ip networks: a survey,” *International Journal of Satellite Communications and Networking*, vol. 21, pp. 303–349, 2003.
210. ETSI, “Satellite earth stations and systems (ses); broadband satellite multimedia (bsm) services and architectures: Qos functional architecture,” vol. TS 102 462, 2006.
211. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, “An architecture for differentiated services,” vol. RFC 2475, 1998.

Index