The EuQoS System: A Solution for QoS Routing in Heterogeneous Networks

X.Masip-Bruin[1], M.Yannuzzi[1], R.Serral-Gràcia[1], J.Domingo-Pascual[1], J.Enríquez-Gabeiras[2], M.Callejo[2], M.Diaz[3], F.Racaru[3], G.Stea[4], E.Mingozzi[4], A.Beben[5], W.Burakowski[5], E.Monteiro[6], L.Cordeiro[6]

[1] Technical University of Catalunya (UPC), Advanced Networks Architectures Lab (CRAAX), Vilanova i G., Catalunya, Spain
[2] Telefónica I+D (TID), Madrid, Spain
[3] Analysis and Architecture of Systems Lab (LAAS), Toulouse, France
[4] Dipartimento di Ingegneria dell'Informazione, University of Pisa, Pisa, Italy
[5] Warsaw University of Technology (WUT), Warsaw, Poland
[6] University of Coimbra (UoC), Coimbra, Portugal

Abstract—**EuQoS is the acronym of "End-to-end Quality of Service support over heterogeneous networks", which is an IST research project aimed at building a complete QoS framework, addressing all the relevant network layers, protocols, and technologies. EuQoS customers will be able to subscribe to the *EuQoS system*, for which we are currently proposing and developing novel QoS mechanisms and protocols which build upon the state of the art. Among these we have: i) Security (AAA); ii) Admission Control; iii) Charging; iv) Signaling and Service Negotiation; v) Monitoring and Measurements; vi) QoS Routing (QoSR); vii) Failure Management; viii) Traffic Engineering and Resource Optimization. The EuQoS system, currently being deployed as a prototype including all the above features, encompasses the most common access networks, i.e., xDSL, UMTS, WiFi, and LAN, which are connected through a core network composed by the National Research and Education Networks (NRENs) of involved partners, and GEANT (the European research network). In this paper we specifically describe the QoS routing mechanisms that are being developed and tested in the frame of the project. We also describe the progress made so far, and the evaluation methodology being used for validating the proposed models and mechanisms. The preliminary performance results validate the design choices of the EuQoS system, and confirm the potential impact that this project is likely to have in the near future.**

Keywords—end-to-end QoS, multidomain heterogenous networks, QoS routing, interdoman routing.

I. Introduction

New demands for using multimedia applications over the current Internet, such as IP telephony, video, tele-medicine, tele-engineering, tele-education, etc, have triggered a spur of research aimed at providing network customers with the required Quality of Service (QoS) [1], in terms of bandwidth, delay, jitter, packet loss, and reliability. Despite routing decisively contributes to the provision of QoS, many factors prevent QoS Routing (QoSR) from being widely deployed. Two of these factors are the most relevant. On the one hand, the problem of QoSR with multiple constraints is known to be NP-hard. This means that while numerous heuristics have already been proposed, only few exact solutions exist [2]. On the other hand, delivering end-to-end QoS to users connected to the Internet through different access networks requires several other building blocks to be properly engineered and interconnected, which is still a big challenge for the research and industry community. Several hot topics, such as admission control, signaling protocols, Traffic Engineering (TE), traffic control, network management, etc, need further research efforts in order to find solutions appealing enough to challenge the usual overprovisioning strategies.

The EuQoS project [3] (an Integrated Project of the 6th IST Framework Program) brings together research centers, universities, Telcos and consulters working in all the abovementioned QoS topics. The main EuQoS target is to define and implement an architectural network model (the *EuQoS system*) capable of guaranteeing end-to-end QoS across heterogeneous networks. EuQoS subscribers should be able to use both registered ("EuQoS-enabled") and legacy applications to communicate with a guaranteed and certifiable QoS. This requires coordinated QoS mechanisms to be placed both in the applications and in the network.

At this stage, the EuQoS team has already designed and developed a first prototype of the EuQoS system (including the main building blocks, such as QoS Routing algorithms, resource allocators, call admission control, signaling mechanisms, etc). This prototype is currently being deployed in a testbed, illustrated in Fig. 1. The latter is built based on GEANT (the European research network) acting as core network, and the NRENs (National Research and Education Networks) connecting different access network technologies such as xDSL, UMTS, WiFi, and LAN.

In this paper, we report an overview of the experiences gained in developing and evaluating such a complete QoS het-

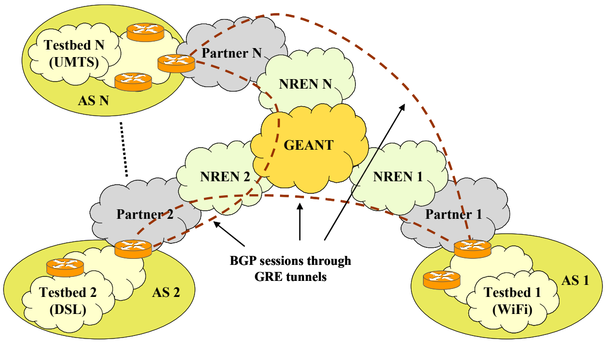


Fig. 1. Access networks, NRENs and GEANT architecture.

This work was partially funded by the European Commission through the EuQoS Integrated Project (FP6-004503) and by the Spanish Ministry of Science and Technology (MCyT) under contract FEDER-TSI2005-07520-C03-02 and the Catalan Research Council (CIRIT) under contract 2005-SGR00481 for UPC's authors.

erogeneous network architecture. More specifically, we concentrate on the network support, with specific focus on QoS routing issues.

The rest of the paper is organized as follows: In Section II the EuQoS architecture and QoS model are presented. In Section III, the EuQoS approach to QoS routing is addressed in detail. A preliminary evaluation of the proposed solutions is shown in Section IV, while Section V concludes the paper highlighting the open issues and the directions for future work.

II. The EuQoS Architecture and QoS Model

The EuQoS network architecture has been defined according to the following high-level rules: i) the customers' applications should be able to negotiate the content and quality of each communication; ii) network administrators should have the freedom to define and use any of the existing network technologies, and the EuQoS system should be deployable on top of them; iii) the proposed mechanisms should be incremental, in the sense that they should coexist with the existing Internet structure.

Fig.2 illustrates the EuQoS architecture. In the rest of this section we overview the main building blocks in Fig. 2, as well as the QoS model, the key signaling mechanisms, and the Monitoring and Measurement System. The QoSR mechanisms are addressed in detail in Section III.

A. Main building blocks

In EuQoS, QoS resource management is handled on a per-session basis. The first step is to derive the QoS requirements from the remote communicating applications. This is accomplished when the communicating applications agree on the content and quality of their communication, i.e. after the application-to-application negotiation. This negotiation requires a signaling protocol between the end-users. In the EuQoS framework, the application signaling is called EQ-SIP, which is an extension of SIP. EQ-SIP includes mechanisms for QoS negotiation of particular QoS parameters (under standardization in the SDPng within the IETF). Therefore, EuQoS application support includes:

- A QoS Control Module (QCM) that links the QoS requests of the users to the network connection.
- Application Signaling (ASIG) that implements the EQ-SDP and EQ-SIP protocol in the users' machines.
- An Extended QoS API (XQoS) that defines the needed different QoS codings.
- A Multicast Middleware protocol that uses a set of point-to-point QoS connections to link them inside an application-level tree-based structure.
- An enhanced transport protocol that provides the QoS adaptations needed to handle the different basic QoS classes in the network layer.

In order to supply the desired freedom to network administrators, a *virtual network layer* has been defined, which decouples network decisions from network technologies. To achieve this goal, this virtual network layer is split into a *technology independent* (*TI*) and a *technology dependent* (*TD*) sub-layer.

AS shown in Fig. 2.a) the TI sub-layer consists of a logical entity, called Resource Manager (RM), which is in charge of managing QoS for each domain. For instance, it coordinates domain-wide admission control decisions, it stores and manages peering agreements with neighboring domains, and it controls the interdomain routing process. Whenever appropriate, the RM decisions are enforced in network devices by means of Resource Allocators (RAs), which are located at the TD sub-layer. Different device technologies have different RAs. The virtual network layer includes the following modules:

- The Signaling and Service Negotiation (SSN) that encompasses support for application signaling, horizontal signaling between the RMs, and vertical signaling between the RMs and the RAs.
- The Connection Admission Control (CAC) in each domain. The CAC module in the RM checks for availability of resources both inside the domain (intradomain CAC) and in the link between peering EuQoS domains (interdomain link

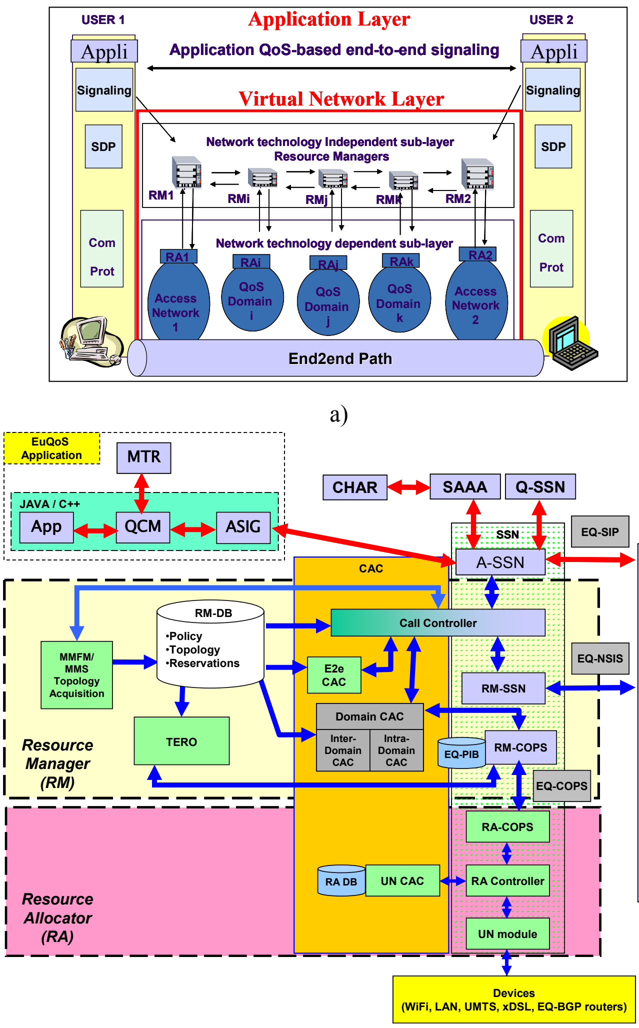


Fig.2. The EuQoS model; a) The high level EuQoS architecture; b) The main building blocks.

CAC). In addition, CAC at the RA level is also enforced.

- The Monitoring and Measurement system (MMS) that provides a dedicated system in order to evaluate the real values of the QoS metrics provided by the network.
- The Traffic Engineering and Resource Optimization block (TERO), which is in charge of interdomain routing configuration and resource provisioning. The TERO actions are described in Section III.
- The Security AAA (SAAA) and Charging (CHAR) modules, which are self-explanatory.

Each application QoS request reaches the virtual network layer through the API (Application Programming Interface) of the access network where the caller application is located. Upon receiving the requests, the TI sub-layer checks the feasibility of an end-to-end path, i.e. the capability of all the networks involved to provide the requested QoS. As a result, an end-to-end path fulfilling the QoS demands needs to be computed. This end-to-end path can be determined following two different approaches:

- Using a "loose" model, in which the data path is determined by a QoSR protocol on a per-Autonomous System basis.
- Using a "hard" model, in which the data path, or part of it, is established by using a traffic engineering mechanism (e.g., MPLS-TE).

In both cases, a signaling protocol is needed for call setup, but the above two approaches are conceptually quite different. In the loose model, all messages (including the signaling messages) will follow the routed data path. On the other hand, in the hard approach, all messages follow pre-specified paths (called *tunnels*), which are built using e.g. MPLS. The two models are not alternative, meaning that the hard model can coexist with the loose model. The loose model is currently implemented in the prototype, and it is the focus of the description of this paper. The hard model is currently being developed by the EuQoS team.

As the EuQoS system is targeted for guaranteed QoS, BGP-4 cannot be used as the interdomain routing protocol. Thus, for the loose model the *Enhanced QoS Border Gateway Protocol (EQ-BGP)* has been developed, building upon a former extension of BGP-4 called qBGP [4]. EQ-BGP is the protocol in charge of determining the QoSR paths between end-users, and it is described in Section III.

As mentioned above, during the call setup phase, the signaling messages will follow the routed data path. However, in order to check whether resources are available, signaling messages must reach each RM along the path. As a consequence, signaling messages have to be forwarded *out* of the normal routing path. To achieve this goal, an extension of the Next Steps in Signaling (NSIS) protocol, currently proposed for standardization at the IETF, has been designed. We define this extension as *EQ-NSIS*. This pioneering implementation of NSIS is used for signaling and exchanging the QoS requirements between the RMs across the different domains (see Fig. 2.a)).

In summary, the end-to-end QoS paths are built using the following key components: i) the set of the RMs; ii) EQ-BGP; iii) EQ-NSIS; and iv) the set of the RAs. The different building blocks of the architecture are illustrated in Fig. 2.b).

B. *The EuQoS QoS model*

Providing end-to-end QoS in EuQoS is achieved by implementing a set of end-to-end Classes of Services (CoSs), as shown in Fig. 3 (following the IETF recommendations [5]). The end-to-end CoSs are known and are visible by the applications (end-users). The traffic generated by a given application is submitted to the appropriate end-to-end CoS once the connection setup process has been successfully completed. Furthermore, all the functions in the RM, the RA, and the EQ-BGP routers are CoS-specific. For instance, different routing tables and routing decision processes exist, different provisioning strategies, traffic control mechanisms, call admission control policies, etc. Each domain is free to provide its own implementation of (a subset of) the defined CoSs, as far as it is compliant with the specifications in Fig. 3. Neighboring domains establish per-CoS peering agreements, called *peering Service Level Specifications* (p-SLSs), which regulate the transit of traffic through the interdomain links.

C. *Signaling in EuQoS*

Two different signaling mechanisms play an important role in order to perform the resource reservations in the EuQoS virtual network layer. These are EQ-NSIS and EQ-COPS. In the sequel, we briefly summarize the main concepts of these two signaling mechanisms.

EQ-NSIS: NSIS is a new protocol being developed in the NSIS working group at the IETF [6]. This working group is responsible for standardizing an IP signaling protocol following a two-layer signaling paradigm with QoS signaling as the first use case. This paradigm consists of a signaling transport layer and the signaling application layer. With this approach the transport of the signaling messages and the signaling application are separate, which allows the protocol to be used for more general purposes. The signaling transport layer is responsible for the exchange of signaling messages among network entities, and it is independent of the signaling applications. The signaling application layer contains the specific functionalities of the signaling applications. This two-layer protocol model allows supporting various signaling applications, including QoS.

In the EuQoS system, as already mentioned, there is the need for involving the RMs in the end-to-end network signal-

End-to-end Class of Servce	QoS Objetives			Type of connections	Traffic descriptor
	IPLR	Mean IPTD	IPDV		
Telephony	10^{-3}	100 ms	50 ms	P2p	Peak rate
RT Interactive	10^{-3}	100 ms	50 ms	P2p	Peak rate
MM Streaming	10^{-3}	1 s	U	P2p	Peak rate
High Thruput Data	10^{-3}	1 s	U	P2p	Requested rate
Standard	U	U	U	-	-

IPLR: IP Packet Loss Ratio ;IPTD: IP Packet Transmission Delay IPDV: IP Packet Delay variation; U: Unspecified; P2p: Peer-to-peer

Fig. 3. Specification of end-to-end CoSs in EuQoS.

ing along a routed data path. Unfortunately, the on-path NSIS protocol is neither able to signal the RMs along the data path, nor to force the signaling messages to follow the same path as the data path [7]. Indeed, the requirements for a hybrid on-path/off-path approach for end-to-end signaling are not fully solved by the NSIS protocol, as it is being currently discussed at the IETF NSIS working group. The EuQoS project team is actively contributing to this work.

The major requirements to achieve successful end-to-end network signaling in the EuQoS system are the following.

- Signaling messages must follow the same path as the data path.
- All the RMs along the data path must be signaled.

In order to fulfill the above requirements, a middle layer between the two NSIS layers has been conceived. This layer is named *Hybrid Path* (HyPath) and was proposed to IETF to be included in the NSIS framework [8]. In order to connect the HyPath with the NSIS Transport Layer Protocol (NTLP) [7] and the NSIS Signaling Layer Protocol (NSLP) [9], without altering the specifications of these latter, the HyPath needs to be a middle layer between the NTLP layer and the NSLP layer. The already defined interfaces between the NTLP and the NSLP remain unchanged. Fig. 4a) illustrates the EQ-NSIS protocol architecture including the HyPath layer. The operation of EQ-NSIS with the additional HyPath layer in the border routers and the RMs in the different domains is illustrated in Fig. 4b).

When a user makes a QoS request to the EuQoS system, EQ-NSIS signaling starts and it must reach all the RMs along the path. This signaling flow must follow the same path as the data path. Therefore, in the first domain, the HyPath in the local RM uses the RM's routing module (the RMs are EQ-BGP-aware) to discover the local border (egress) router for the data path. After that, the HyPath asks the NSIS transport layer to send a NSIS message to the corresponding border router. This message contains the NSLP payload and some additional HyPath information. Once in the border router, the EQ-NSIS signaling message is sent toward the end-user's domain. In this scenario, all border routers are HyPath aware. In each downstream domain the EQ-NSIS signaling message is intercepted by the ingress border router and redirected to the local RM (see Fig.4b)).

After processing the message, each RM resumes the signaling sending a message back to the ingress border router. The signaling is restarted in the ingress border router and the NSIS message continues toward the next domain. This process continues along all downstream domains until the last domain is reached. With this architecture all the requirements to achieve end-to-end network signaling are met and no changes are needed in the definitions of the NTLP and NSLP layers.

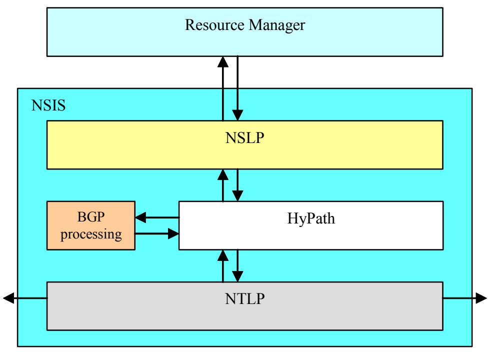
EQ-COPS: This signalling protocol is an extension of the Common Open Policy Service (COPS) [10] developed by the EuQoS team. EQ-COPS is utilized for vertical signaling between the RMs and the RAs, and it plays an important role in the EuQoS framework. It provides a technology independent scheme to map high-level QoS domain policies into low-level network device configurations, coping both with the required autonomy of QoS management inside each domain and the need to establish a TI sub-layer composed by the RMs.

Each administrative QoS domain maintains its own Policy Repository (PR) and its own Policy Decision Point (PDP). The PR stores domain-specific policies according to a Lightweight Directory Access Protocol (LDAP) scheme. Those policies are then used at two distinct levels:

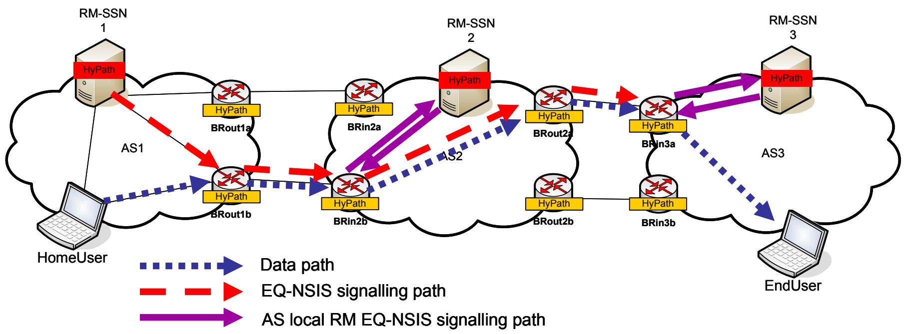
- To define the technology-independent behavior of the RMs, i.e., which QoS requests should be satisfied and under what circumstances.
- To translate the technology-independent QoS requirements into specific network device configurations, using EQ-COPS-PR for the communication between the PDP (located in the RM) and the Policy Enforcement Points (PEPs) which are located in the RAs.

D. The Monitoring and Measurement System (MMS)

The MMS is a subsystem designed for performing specific



a)



b)

Fig. 4. NSIS architecture: a) HyPath in the EQ-NSIS architecture; b) EQ-NSIS signaling with HyPath.

measurements of the network traffic and to check the status and availability of network resources. The MMS is utilized to: i) evaluate and confirm that the network fulfills the QoS demands of the ongoing sessions in the EuQoS system; ii) assess the availability and performance of the network resources in the EuQoS system.

Fig. 5 shows the structure of the MMS subsystem and its interface with the *Measurement, Monitoring and Fault Management* (MMFM) module, which is actually part of the RM. The MMFM is the higher level entity in charge of storing the network data provided by the MMS. The network information is stored in the RM Database. In addition, the MMFM module delivers these data to some other specific modules such as the TERO and the CAC when needed (for instance in the event of a node or a link failure). The MMS on the other hand, is the lower level entity in charge of providing the network information. With the aim of guaranteeing the QoS for the ongoing sessions, the information is provided in real-time to the MMFM. This allows the corresponding actions to be triggered when an event affects the delivered QoS.

The captured data stores the usual quality of service parameters, such as One-Way Delay, Packet Losses, Jitter, etc, the subjective quality parameters, like the Mean Opinion Score (MOS) for VoIP, the current link load and the interdomain topology, which are used to control routing decisions.

For capturing the required network traffic, the MMS needs several capture points, specifically, in the ingress point of the access network and on the egress (as shown in Fig. 5). This, using flow detection mechanisms [11], enables MMS to compute QoS parameters per-CoS on an intradomain basis.

In interdomain and end-to-end QoS validation situations, the higher MMFM layer must share the information within the MMFM sub-system. This is accomplished by selecting the appropriate capture points and by sharing the flow information from the database where it is stored.

The most critical point of this sub system is to react to unexpected network events with the minimum possible delay. For such efficient reporting to the MMFM, a structured interface has been developed between the MMS and the MMFM. This interface has three main messages: i) *Result Request*, which allows the MMFM to request any available results from MMS; ii) *Changes Notification* which allows one to react to unexpected situations (i.e. topology changes); iii) *Configuration*, which changes the MMS behavior regarding the traffic being monitored or the way the statistics are reported.

III. QoSR IN THE EUQOS SYSTEM

The EuQoS system targets to provide end-to-end QoS paths over heterogeneous networks. This motivates to encompass routing issues concerning QoS at the access networks, at the intradomain, and at the interdomain level. So far, the major research efforts in the project have been devoted to interdomain QoSR, which is considered the most important issue by the telecom operators participating in the project. Furthermore, in most practical settings the users' terminals (UMTS mobile phone, a WiFi notebook, a DSL modem/router, etc) are typically connected through stub networks. Thus, routing in these

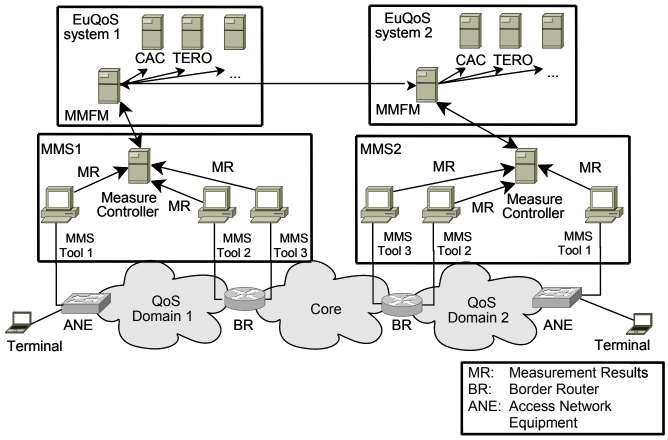


Fig.5. The Monitoring and Measurement System (MMS).

terminals is usually handled by means of default routing. This means that in practice the QoSR decisions need to be made *between* the source and destination access networks, but not *within* the latter.

It is worth highlighting that even though QoSR decisions are not necessarily needed inside the access networks, QoS still needs to be delivered in these networks. In the EuQoS system this is managed by means of CoS subscription, QoS policies (CAC on a per-CoS basis and traffic shaping), resource reservations during an EuQoS session, and monitoring. For instance, the approach followed to integrate UMTS access networks into the EuQoS system lies on using the built-in mechanisms provided by the UMTS technology to set-up and route connections (PDP contexts) [12]. In a WiFi network, separate access points are provided for QoS and best effort traffic, which therefore do not interfere with each other. LAN and xDSL access networks are similarly handled. The proposal for these networks is the replacement of the Subnet Bandwidth Management (SBM) solution proposed by the IETF, with a solution based on the use of NSIS signaling for resource allocation in Ethernet and xDSL networks. In this proposal, the RM is the entity responsible for the admission control in the Ethernet/xDSL domain, having similar functionality to the Designated SBM defined by IETF.

There are many new mechanisms implemented within the EuQoS project related to QoSR issues. Hereafter we describe the EQ-BGP protocol, the Topology Acquisition Tool (TAT), and the TERO module, as applied to interdomain QoSR.

A. Interdomain QoSR

In the interdomain QoSR process of the EuQoS system, an AS negotiates *peering Service Level Specifications* (p-SLSs) with its neighbors. P-SLSs regulate the transit and QoS guarantees of traffic belonging to a given CoS at an interdomain link, in one direction. Thus, two ASs negotiating a p-SLS are called the *customer* and *provider* of that p-SLS, meaning that the traffic flows from the former to the latter. More specifically, an AS sends EQ-BGP advertisements for a given CoS *only* along interdomain links at which it has previously negotiated a p-SLS as a provider. P-SLSs formally specify:

- The amount of traffic that a customer can inject at the interdomain link, and the actions that the provider will take against non-conforming traffic.
- The QoS that the provider guarantees to the admitted traffic.

Packets of a given CoS can leave an AS through an interdomain link only if a p-SLS exists for that CoS at that interdomain link. Thus, interdomain QoSR is constrained by p-SLSs, which are controlled by the TERO module. In the sequel we describe the main components related to interdomain QoSR in the EuQoS system.

EQ-BGP: The EQ-BGP protocol was developed within the EuQoS project with the aim of performing interdomain QoSR. The objectives of EQ-BGP are to advertise and select the routing paths for the different CoSs in Fig.3. EQ-BGP extends the BGP-4 routing protocol in the following way. First, EQ-BGP includes an optional path attribute, named QoS_NLRI that conveys information about the QoS capabilities of a path. Second, it includes a QoS assembling function for computing aggregated values of the QoS parameters for the whole routing path. In general terms, this assembling function can supply the sum of the delays for each segment of a path, or the minimum available bandwidth along a path. Third, EQ-BGP has a QoS-aware decision process for selecting the best end-to-end path for the different CoSs. And fourth, EQ-BGP handles multiple routing tables in order to store the available paths for each end-to-end CoS.

EQ-BGP performs QoS routing in multidomain networks by taking into account both intra- and interdomain QoS information. For that purpose, each EQ-BGP router advertises to its neighbors the reachable destination addresses including information about the available end-to-end CoSs. On that basis, each EQ-BGP router selects the best QoS path for each end-to-end CoS and informs its neighbors about its choice. Thus, EQ-BGP sets the road-map for the available QoS paths between each pair of source and destination networks. These paths are called *end-to-end QoS paths* and they are computed and advertised by EQ-BGP routers for each CoS separately.

In Fig. 6 we show an example of how the QoS routing information is computed and advertised across different domains using EQ-BGP. For the sake of simplicity, we assume a simple network consisting of three domains A, B and C that support the same end-to-end CoSs. We assume that each EQ-BGP router is aware of the *nominal* values of the QoS parameters that are assured both inside its particular domain (Q_A, Q_B, or Q_C depending on the domain) as well as on its corresponding interdomain links ($Q_{A\text{->}B}$, $Q_{B\text{->}A}$, $Q_{B\text{->}C}$, or $Q_{C\text{->}B}$ depending also on the domain). It is important to highlight that all these nominal QoS values are computed by the TERO module during the network provisioning process and they correspond to the *Maximum Admissible Load* controlled by the intra- and the interdomain call admission control functions. The reason for this is to avoid route flapping due to frequent variations of the QoS values. These nominal QoS parameters typically change at provisioning timescales (e.g., in the order of days, or weeks). This approach provides a scalable EQ-BGP routing protocol, but clearly the success of the approach requires adaptive provisioning and strict admission control policies. During our simulations, the approach of using and advertising aggregated nominal QoS parameters in conjunction with adaptive provisioning has proven to achieve excellent results.

Now, let us consider the case when domain C advertises a new prefix, say $NLRI_C$. Then, the routing information is propagated toward domain A through domain B. Fig. 6 shows how the QoS routing tables of the border EQ-BGP routers become populated along the path. During this process EQ-BGP routers aggregate the nominal values of the QoS parameters along the path taking into account the nominal QoS contributions of the intradomain segments as well as those of the interdomain segments of the path. For example, domain A learns an end-to-end QoS path toward the destination $NLRI_C$, with QoS corresponding to $[Q_C \oplus Q_{B\text{->}C} \oplus Q_B \oplus Q_{A\text{->}B}]$ for a particular CoS, wherein the operator $\oplus$ denotes an appropriate QoS assembling function. The EQ-BGP decision process is described in the next subsections.

TAT: The information about the network state is obtained by the Topology Acquisition Tool (TAT). TAT is aimed at collecting information from the routers about both the currently used and the alternative interdomain paths that are set by the EQ-BGP protocol.

TERO: The EuQoS system architecture also includes the interdomain Traffic Engineering and Resource Optimization (TERO) module which is located in the RM. TERO is in charge of interdomain routing configuration and resource provisioning. More specifically, it controls the interdomain routing process, so as to steer the traffic through the ASs in the most effective way, optimizing interdomain resources (i.e., bandwidth and buffer space on the interdomain links) based on QoS requirements. Furthermore, it configures queues and policers at interdomain links so as to provision the necessary resources to allow traffic to flow across neighboring domains. TERO actions regarding routing and provisioning can be taken either as a reaction to the variation of the network topol-

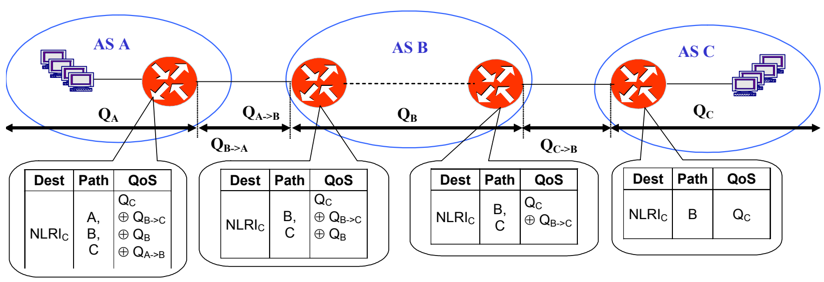


Fig.6. Example of EQ-BGP operation.

ogy or periodically, for maintenance and optimization. Thus, TERO works at a *network provisioning* timescale, (i.e. days or weeks), i.e. at a much larger timescale than an EuQoS session lifetime.

TERO interacts with the border routers through EQ-COPS, so as to configure the EQ-BGP protocol. More specifically, TERO configures EQ-BGP routers so that:

- EQ-BGP UPDATE messages are allowed to flow through an interdomain link whenever a new p-SLS is negotiated.
- The QoS-NLRI information is properly updated before EQ-BGP messages are advertised to neighboring ASs as well as inside the domain. In fact, the QoS-NLRI information advertised to upstream domains assembles the *nominal* QoS-NLRI information included in the UPDATE messages received from downstream domains, with the *nominal* values of the QoS parameters that are assured both intra- and interdomain links of the domain (see Fig. 6).
- When an EQ-BGP router receives multiple updates for the same destination, it runs the EQ-BGP decision process. This process includes an additional step when compared with the BGP-4 decision process, which takes into account a new parameter called the *Degree of Preference* (*DoP*). The latter is computed by EQ-BGP routers based on the *QoS preference parameters* provided by TERO. These QoS preference parameters are defined in subsection *A.2*.

Hereafter, we provide further details on these configuration actions.

A.1. EQ-BGP configuration process

The EQ-BGP configuration process takes place when a p-SLS is added or removed at an AS. In fact, the activation (termination) of a p-SLS at an interdomain link, besides starting (stopping) the flow of traffic from the customer to the provider, also starts (stops) the propagation of EQ-BGP updates in the opposite direction. Assume a new p-SLS is activated at the A-B interdomain link of Fig. 6. The AS A has to be provided with the QoS preference parameters (defined in the next subsection) for the negotiated CoSs. Furthermore, the border router in AS A has to be provided with the contribution to the QoS parameters of the interdomain link ($Q_{A\text{->}B}$), with which to update the incoming updates before propagating them inside AS A. At AS B, instead, the contribution to the QoS parameters due to intradomain traversal has to be configured on the ingress router, so that it latter can “add” (⊕) Q_B to the updates flowing from B to A. Note that, since new traffic is now injected into AS B, the contribution of intradomain traversal has to be recomputed also for other border routers (not shown in the figure) at which a provider p-SLS is in place. When a p-SLS is terminated, similar actions occur. P-SLS management is implemented in EuQoS through a simple web interface.

A.2. Degree of Preference (*DoP*)

When a border router receives an EQ-BGP update from another AS, it associates a DoP to that update. The latter is exploited in the EQ-BGP decision process to select one among different updates advertising the same destination for a given CoS. The DoP parameter has a *local*, AS-wide meaning, and it is never advertised to other ASs.

Since in EuQoS there is one decision process for each CoS, the computation of the DoP can - in principle - be different from one CoS to another. In the EuQoS prototype, the following formula has been used for all CoSs in Fig.3, except the Standard one:

$$DoP = \sum_{i \in \{IPTD, IPDV, IPLR\}} \frac{f_i}{\max\left\{0, \left[M_i - \left(Q_i \oplus Q_i^{'}\right)\right]\right\}} \qquad (1)$$

The DoP computed in an EQ-BGP border router is then the sum of three terms, each one associated to a different QoS parameter carried in the EQ-BGP updates (i.e., IPTD, IPDV, and IPLR). Each term consists of:

- A *QoS Preference* f_i, which accounts for the relative importance of the QoS parameter i with respect to the others.
- A *parameter value* Q_i, i.e. the assembled value of the QoS parameter i carried in the incoming update.
- A *parameter value* Q_i’, i.e. the actual (real) value of the QoS parameter i in the interdomain link from which the EQ-BGP router receives the incoming update. This parameter basically takes into account the current load on the interdomain links and it is used to *locally* compute (1), but it is never included in the QoS-NLRI information advertised to upstream domains (for scalability reasons only nominal QoS values are assembled and advertised to other domains).
- A *maximum value* M_i allowed for the QoS parameter i, taken from ITU recommendations.

If a border router receives more than one update for the same destination, it selects the one with the *lowest* DoP. In fact, the DoP increases with the value of the QoS parameters, and it goes to infinite (forcing the decision process not to select a specific route) if the value of one of the assembled QoS parameters exceeds the maximum M_i. Fig. 7 summarizes the EQ-BGP decision process.

Different QoS preferences f_i are assigned by TERO to different CoSs. For instance, for Telephony and Real Time Interactive the IPTD and IPDV are equally important and more important than IPLR, whereas for Multimedia Streaming IPTD is more important than IPDV, and for High Throughput Data IPLR is the most important. While the above mechanism provides a sufficient degree of flexibility, fine tuning of the QoS preference parameters requires extensive simulations and tests on the EuQoS prototype.

1. Choose the route with the highest local preference (LOCAL_PREF)
2. If the LOCAL_PREFs are equal choose the route with the lowest Degree of Preference (DoP)
3. If the DoPs are equal choose the route with the shortest AS-path
4. If the AS-path lengths are equal choose the route with the lowest MED
5. If the MEDs are equal prefer external routes over internal routes (eBGP over iBGP)
6. If the routes are still equal prefer the one with the lowest IGP metric to the next-hop router
7. If more than one route is still available run tie-breaking rules

Fig. 7. EQ-BGP decision process.

IV. Evaluation Methodology

In this section we evaluate the performance of EQ-BGP. Our objective is to analyze the impact of the new components in EQ-BGP on its scalability. The evaluation is performed by comparing the performance of EQ-BGP against BGP-4 in different network scenarios. In order to perform this comparison we assess two different metrics, namely:

- **The Network Convergence Time** (NCT) defined as the total amount of time that elapses between the advertisement of a new prefix (or the withdrawal of known one), and the time instant when the last update message caused by this event is processed.
- **The total Number of Messages** that are exchanged during the network convergence time.

In our experiments we consider three types of network topologies with a different number of ASs, i.e., full-mesh, ring and a representative topology for the Internet. The full-mesh topology was selected because it allows to have the maximum number of alternative paths, thus representing a "worst case" scenario. On the other hand, the ring network (or b-clique) is commonly used for analyzing the routing decision algorithm, as there are exactly two disjoint paths between each pair of domains. To complete the performance evaluation, we analyze the performance of EQ-BGP in topologies derived from operating networks as presented in [13]. This network model, called "Internet" was derived from routers operating in the Internet backbone.

For the sake of simplicity, we assume that each AS is represented by a single EQ-BGP router connected with its neighbors using links of 1Mbps of capacity, introducing a constant delay of 1msec. Although the link parameters were arbitrarily chosen, the results obtained for the NCT can be easily scaled taking into account actual link characteristics. In addition, we assume that all the ASs support the Telephony CoS, guaranteeing different values of IPTD. The values of IPDV, IPLR as well as the corresponding parameters on inter-domain links, are the same for all domains. Moreover, we consider three different strategies for the assignment of IPTD values to specific domains that are: i) random strategy, i.e. the value of the delay offered by a given AS is randomly chosen from 1ms to 10 ms; ii) increasing strategy, i.e. the delay increases with the AS number; and iii) decreasing strategy, i.e. the delay decreases as the AS number increases (in other words, ASs with lower AS numbers introduce larger delays).

Our experiments were performed using the ns2 simulator [14], in which the EQ-BGP protocol has been implemented. All experiments were performed assuming that the advertisement or withdrawal of a prefix occurs when the network is in a stable state (i.e. after it has already converged). Each simulation run was stopped when the last update message originated by the considered stressing event was processed. The results presented here were collected from 10 simulation runs, in which a randomly chosen AS advertises or withdraws a route. The reported values of convergence time include the 95% confidence interval. The next subsections present the results obtained both in terms of the network convergence time and the number of messages exchanged during convergence.

A. *Network Convergence Time*

In Fig. 8.a) we present the results corresponding to the NCT of "full-mesh", "ring" and "Internet" netwrok topologies, after the advertisement of a new route. The results obtained show that the full-mesh network exhibits the same convergence time for both EQ-BGP and BGP-4, irrespective of the number of ASs. This can be explained by considering that all the ASs select their routing paths on the direct links. As the QoS level assured on direct links is usually better than the one offered on the alternative paths, none of the routers will switch its routing to an alternative path. Thus, the routing process ends at the same time for all the cases assesed. On the other hand, for the

Network type	No. of ASs	BGP-4		EQ-BGP					
				random		increasing		decreasing	
		mean [ms]	max [ms]	mean [ms]	max [ms]	mean [ms]	max [ms]	mean [ms]	max [ms]
"Full mesh"	4	3.55±0,0	3.55	3.55±0,0	3.55	3.55±0,0	3.55	3.55±0,0	3.55
	11	3.55±0,0	3.55	3.55±0,0	3.55	3.55±0,0	3.55	3.55±0,0	3.55
	29	3.55±0,0	3.55	3.55±0,0	3.55	3.55±0,0	3.55	3.55±0,0	3.55
"Ring"	4	6,5±0,028	8,22	6,78±0,028	8,22	7,64±0,023	8,22	6,79±0,028	8,22
	11	10,84±0,0	10,85	12,94±0,028	14,6	13,32±0,025	14,6	13,49±0,027	14,6
	29	28,2±0,0	28,2	51,45±0,045	53,94	47,26±0,05	50,4	47,79±0,06	50,4
"Internet"	29	11,92±0,028	15,31	17,9±0,076	27,38	13,59±0,06	20,38	18,97±0,065	27,36

a)

Network type	No. of ASs	BGP-4		EQ-BGP					
				random		increasing		decreasing	
		mean [ms]	max [ms]	mean [ms]	max [ms]	mean [ms]	max [ms]	mean [ms]	max [ms]
"Full mesh"	4	16,75±0,03	18,47	16,83±0,03	18,47	17,04±0,03	18,47	12,49±0,0	12,49
	11	1854±6,13	2077	717,2±4,21	1139	1275±3,5	1544	576,5±1,7	708,8
"Ring"	4	8,22±0,03	9,56	7,89±0,03	9,56	6,89±0,02	9,56	7,89±0,03	9,56
	11	20,03±0,0	20,03	17,84±0,03	20,03	17,52±0,02	20,03	17,52±0,02	20,03
	29	53,94±0,0	53,94	51,45±0,05	53,94	47,26±0,05	50,4	47,79±0,06	50,4
"Internet"	29	1872±13	2808	1212±65	1706	2654±23	4042	2328±26	5241

b)

Fig. 8. Comparison of EQ-BGP and BGP-4 convergence time: a) After a route advertisement; b) After a route withdrawal.

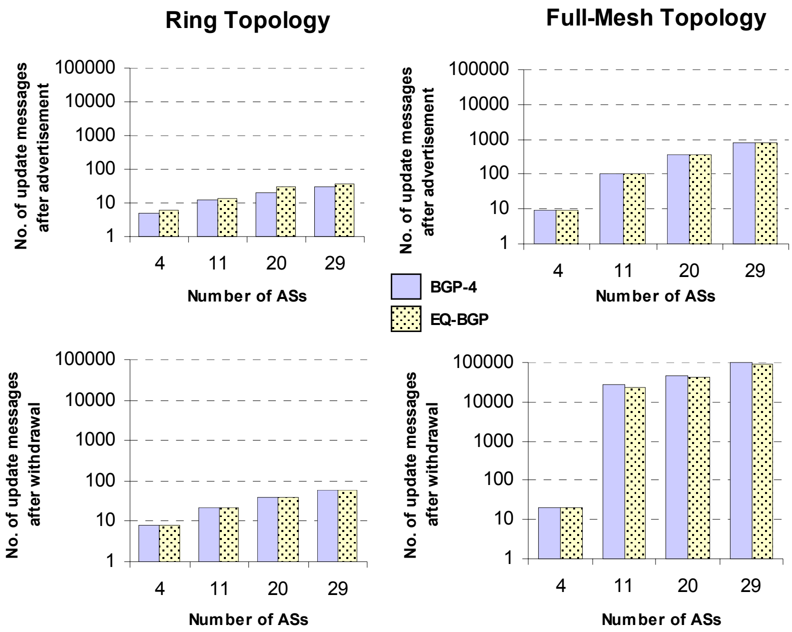


Fig.9 Scalability of EQ-BGP vs. BGP-4

"ring" and "Internet" networks, the NCT time increases with both the number of ASs and the number of interdomain links. Moreover, we can observe that the EQ-BGP protocol needs more time to converge, almost twice as much in some cases. This effect is caused by the introduction of an additional degree of freedom compared to BGP-4, stemming from the possibility of assigning arbitrary QoS parameters in an AS, instead of a single value as in the case of the AS-path length. Therefore, the possibility that an alternative path is better than the one currently used becomes larger. As a consequence, the network exhibits a slower convergence. The opposite effect can be observed in the case of a route withdrawal, as shown in Fig. 8b). For all the types of networks assessed, EQ-BGP usually converges slightly faster than BGP-4. This is because alternative paths have assigned more information about their capabilities, and hence less suitable paths are removed faster. However, such a reduction in the convergence time is in fact negligible.

Within the limits of the preliminary evaluations performed so far, the EQ-BGP protocol has proven to be stable and to exhibit a convergence time comparable to that of BGP-4.

B. *Number of Messages Exchanged during Convergence*

The EQ-BGP protocol is designed to operate in rather large networks. Therefore, assessing its scalability is an important part of the performance evaluation. To accomplish this, we compare the number of update messages processed by both EQ-BGP and BGP-4 during a network convergence, i.e. after the advertisement or withdrawal of a prefix. From the results shown in Fig. 9 we can observe that EQ-BGP and BGP-4 require a similar number of messages to converge.

V. Discussion and Future Work

This paper deals with the problems of finding and providing an end-to-end QoS path between two users connected through heterogeneous access network technologies. This is the main target of the EuQoS research project. In the current state of the project a first prototype has been designed, developed and it is being implemented on a real testbed made up of GEANT, the NRENs (per country) and including different access network technologies (in particular, WiFi, LAN, xDSL and UMTS). This prototype includes the implementation of several specific solutions addressing key points such as signaling and routing (mainly focusing on interdomain at this stage). After evaluating by simulation the first prototype some conclusions and open issues arise. The first and probably the most important, is about scalability issues. The EuQoS system was devised as a solution to provide QoS among a set of peering ASs, but it is not expected that the EuQoS system can become deployed on a wide scale in the Internet. The project team is currently working on this topic, mainly focusing on the signaling mechanisms and the EQ-BGP protocol. Second, new access networks are to be included in the list (G/MPLS and Satellite). Third, modules should be designed for an easier integration. This is an ongoing work grouping all partners involved in the real implementation of the system. And fourth, we are also checking the architecture designed during the first phase of the project with the purpose of devising new, more advanced QoS solutions.

Acknowledgments

We would specially thank all the partners involved in the project for their support, as well as for their work on developing the EuQoS system.

REFERENCES

[1] X.Masip-Bruin, M.Yannuzzi, J.Domingo-Pascual, A.Fonte, M.Curado, E.Monteiro, F.Kuipers, P.Van Mieghem, S.Avallone, G.Ventre, P.Aranda-Gutiérrez, M.Hollick, R.Steinmetz, L.Iannone, K.Salamatian, “Research Challenges in QoS Routing,” Computer Communications Journal, Elsevier, vol. 29, issue 5, pp. 563-581, March 2006.

[2] P.Van Mieghem, F.A.Kuipers, T.Korkmaz, M.Krunz, M.Curado, E.Monteiro, X.Masip-Bruin, J.Solé-Pareta, S.Sánchez-López, “Quality of Service Routing,” Chpater 2 book "QUALITY OF FUTURE INTERNET SERVICES: COST 263 FINAL REPORT", Ed. Springer-Verlag, October 2003.

[3] End-to-end Quality of service support over heterogeneous networks (EuQoS), FP6-IP 004503. http://www.euqos.org/.

[4] G.Cristallo, C.Jaquenet, “An Approach to Interdomain Traffic Engineering,” XVIII World Telecommunications Congress (WTC2002), Paris, France, September 2002.

[5] D Kwok Ho Chan, Jozef Babiarz, Inter-Provider Service Classes, Inter-provider QoS Meeting#2, 28 January 2005, US, http://cfp.mit.edu/qos/ slides.html.

[6] Next Steps in Signaling (NSIS), IETF WG: http://www.ietf.org/html. charters/nsis-charter.html.

[7] H.Schulzrinne, R.Hancock, “GIST: General Internet Signalling Transport,” draft-ietf-nsis-ntlp-09 (work in progress), IETF, March 2006.

[8] L.Cordeiro, M.Curado, E.Monteiro, F.Racaru, M.Diaz, “Hybrid on-path off-path approach for end-to end signalling accross NSIS domains (HyPath),” Internet Draft draft-cordeiro-nsis-hypath-00, IETF, February 2006.

[9] J.Manner, G.Karagiannis, A.McDonald, S.Bosch, “NSLP for Quality-of-Service signaling,” draft-ietf-nsis-qos-nslp-08 (work in progress), IETF, April 2006.

[10] D.Durham, J.Boyle, R.Cohen, S.Herzog, R.Rajan, A.Sastry, “The COPS (Common Open Policy Service) Protocol,” RFC 2748, January 2002.

[11] T.Zseby, S.Zander, G.Carle, “Evaluation of Building Blocks for Passive One Way Measurements”, Passive and Active Measurements Conference (PAM), Amsterdam, Netherlands, April 2001.

[12] 3GPP TS 27.060 - 3rd Generation Partnership Project; Technical Specification Group Core Network; Packet Domain; Mobile Station (MS) supporting Packet Switched Services (Release 1999).

[13] B.Premore, “Multi-AS topologies from BGP routing tables,” http://www. ssfnet.org/Exchange/gallery/asgraph/index.html, 2005.

[14] The Network Simulator 2 (ns-2). http://www.isi.edu/nsnam/ns/.