Architecture and Protocols for the Seamless and Integrated Next Generation IP Networks

Gino Carrozzo[1], Nicola Ciulli[1], Stefano Giordano[2],
Giodi Giorgi[1], Marco Listanti[3], Ugo Monaco[3],
Fabio Mustacchio[2], Gregorio Procissi[2], and Fabio Ricciato[3]

[1] Divisione Informatica e Telecomunicazioni of Consorzio Pisa Ricerche, Italy
{g.carrozzo,n.ciulli,g.giorgi}@cpr.it
[2] Dipartimento di Ingegneria dell'Informazione of University of Pisa, Italy
{s.giordano,fabio.mustacchio,g.procissi}@iet.unipi.it
[3] INFO-COM Dept. of University "La Sapienza" of Rome, Italy
{listanti,monaco,ricciato}@infocom.uniroma1.it

Abstract. The paper presents a novel end-to-end seamless framework to support end-to-end Quality of Service and Traffic Engineering. The network model is based on the MPLS/DiffServ paradigm and addresses the definition of a network architecture according to both users and network providers requirements. A first solution relies on the centralized MPLS/DiffServ based Multi-protocol Access Inter-Domain (MAID) architecture. This architecture allows a seamless QoS-IP service setup through proper Users-Network Interfaces and inter-domain communication through Network-to-Network Interfaces. A fully distributed solution is also presented to address critical scalability issues and to improve network resilience. The overall architecture has been validated by means of functional tests carried out on operational testbeds based on Linux PC platforms.

1 Introduction

The Quality of Service for IP packet flows (QoS-IP) has a long history of standards and tools, both at the Data Plane level (e.g. traffic conditioning) and at the Control Plane level (e.g. signaling and policy protocols). Quality of Service requirements strongly depend on the side in which the interaction users–network is observed. From the user perspective, the basic QoS requirements are the dynamism (e.g. the service should last as long as the user needs), the tailoring (e.g. the network resources allocated for the service should fulfill exactly the end-user requirements), as well as a seamless integration (e.g. the mechanisms involved in QoS support should be transparent to end-user applications). Though some tools for QoS are available in commercial IP routers, their compliancy to these requirements is still far from being a market reality. Indeed, the main obstacles for such a deployment reside in the different backbone networks technologies (e.g. DiffServ, MPLS, IPoATM, etc.), which make hard to guarantee end-to-end QoS, above all when the service has to be deployed across different administrative

M. Ajmone Marsan et al. (Eds.): QoS-IP 2005, LNCS 3375, pp. 419–432, 2005.

domains and in the number of protocols used in the access networks (e.g. RSVP, H.323, SIP, MPEG-4,..), which implies a per-service/per-protocol User-Network-Interface (UNI). From the Service Provider perspective, other requirements drive the evolution of the services offered by the backbone, such as:

- network scalability, which implies a distributed Control Plane, best fitted if based on MPLS;
- traffic engineering both at the flow and at the resource level, best fitted if based on a DiffServ data plane;
- service survivability in case of faults or dynamic network topology changes, easily guaranteed by MPLS recovery strategies;
- interoperation of adjacent domains with the same or different technologies, which implies a Network-to-Network-Interface (NNI);
- interoperation of equipments from different vendors.

The overall objective of the research activity within the TANGO project [1] is to define a novel network architecture able to meet the above summarized users/Service Providers requirements. The crucial point of this architecture will be the definition of proper interfaces between users and network (UNIs) and between network and network (NNIs), respectively.

The paper is organized as follows. Section II presents the paradigm of nested-networks and addresses the concepts of UNIs and NNIs that will be elaborated upon in the next sections. Section III is devoted to the seamless QoS-IP service setup, and includes the definition of the *Multi-protocol Access Inter Domain* (MAID) architecture and of the corresponding UNI (single MAID domain) and NNI (multiple MAID domains). Section IV discuss the key aspects of a MPLS/Diffserv backbone architecture in this scenario. An alternative solution to address critical scalability issues and network resilience is presented. Section V presents the results of functional tests carried out on the experimental test-bed developed in the framework of TANGO project. Finally, Section VI concludes the paper with final remarks.

2 The Nested-Networks Paradigm

The currently operational networks feature a mature IP Data Plane, in which QoS-IP network services are statically configured (and, consequently, under- or over-provisioned) through the Management Plane. These networks are provided with a flat Network Interface (NI) hierarchy, thus in most cases the different NI functions, such as policing and traffic conditioning, are summed up in a single point (the accessing router) even if the network service traverses multiple operators/providers; an example of this single-point Service Level Agreement (SLA) is the Acceptable User Policy (AUP) agreement at the NRENs User-Network Interface (UNI). For these networks, the IETF DiffServ architecture [2] has been largely recognized as a main technology component for QoS-IP networks, due to its native scalability and flexibility.

The DiffServ specifies only mechanisms for packet forwarding (Per Hop Behavior - PHB), flow aggregation rules and traffic conditioning, with a strict Data

Plane scope [3–5]. Internet Service Providers (ISP) configure with internal policies the desired intra-domain services from PHBs, in order to fulfill the Service Level Agreements (SLA) drawn up with their accessing users.

However, although the DiffServ architecture solves, at the Data Plane level, the scalability problems related to QoS provisioning in a single domain, it is not a complete end-to-end solution for the enforcement of a globally dynamic and multi-level chain of SLAs. Different combined solutions (e.g. IntServ/DiffServ, MPLS/DiffServ) have been proposed to overcome the lack of Control Plane procedures and several research projects have been carried out to address this problem (e.g. IST AQUILA [6], IST TEQUILA [7], IST MESCAL [8], IST MOICANE [9], etc.). In general, in those projects where the focus was on intra-domain Control Plane, the inter-domain was out of scope, and vice versa; and this is a major lack for the assessment of an end-to-end seamless framework.

In this scenario, manufacturers can provide network operators with sets of tools (e.g. Bandwidth Broker-like) for managing the QoS parameters of their own network elements. These tools rely on the a common management paradigm, based on standard or, more frequently, on proprietary Management Information Bases (MIBs). However, these tools cannot prove to be effective in multi-region signaling-integrated scenarios, due to the lack of generalized interfaces at the different boundaries of the network (e.g. UNI, NNI), capable of integrating heterogeneous protocols from the access network (e.g. MPEG-4 DMIF, RTP, RSVP or SIP) towards the intra-domain (e.g. RSVP-TE, COPS, SIBS) and, in case, towards the inter-domain. In the following section an architecture aimed at solving the critical seamless interoperabilty issue is presented, with focus on control plane mechanisms and interfaces among different network segments.

3 The Seamless QoS-IP Service Setup

3.1 The MAID Architecture

In current operative IP networks there is a lack of seamless procedures for QoS-IP service setup. The Multi-protocol Access Inter-Domain (MAID) architecture is aimed at providing Network Operators with the robust and user-friendly mechanisms to support QoS in MPLS/DiffServ networks, hiding the underlying complexity of managing all the involved parameters. Concerning the Data Plane this architecture provides the mapping and the forwarding of the IP flows from the access network into the proper DiffServ Label Switched Paths (LSP). On the other hand, the MAID Control Plane is responsible for Admission Control (AC) and policy decisions (taken on a per-flow or per-PHB basis) and for LSPs management. The key elements of the MAID network are the accessing border router (MA-BR), which triggers the setup of QoS-IP services upon receiving QoS requests from the access networks, and the Bandwidth Broker (BB), which manages network resources and policies, as well as inter-domain communications (Fig. 1). The main functionality of the MA-BR is to provide the inter-working between the access network and the backbone. Therefore, the MA-BR manages

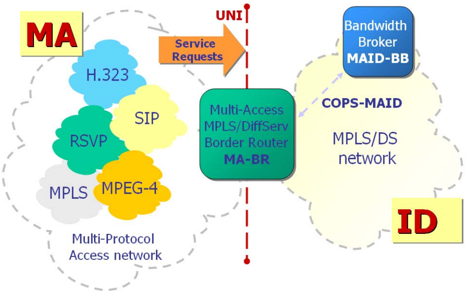


Fig. 1. Multiple Access Inter Domain (MAID) network model.

all the service requests from the access network, defined in terms of protocol-specific QoS semantics, and it conveys them in a generalized and unified service request (i.e. the UNI request) with a unique QoS syntax. If some resources have been provisioned by BB for MA-BR, the processing of the UNI requests for QoS-IP network services (i.e. the AC and policy decisions) can be handled locally to the MA-BR; otherwise, these requests are directly processed by the BB, according to the outsourcing operation model. In any case, BB has pre-emption rights on each MA-BR decision, since it provides a centralized AC and policy that is supposed to be optimal with respect to the local AC provided by MA-BR.

The signaling protocol for the communication between an access network and the MA-BR is application-dependent (e.g. RSVP for IntServ networks, H.323 or SIP for VoIP, DMIF signaling protocol for MPEG-4 services, etc.). Instead, the communications between the MA-BR and the BB are based on a extended version of the Common Open Policy Service protocol (COPS), detailed in the next section. It is possible that BB configures directly core and border routers according to its criteria (e.g. via SNMP protocol). In order to make the entire system scalable, it is desirable to tune an optimum mix of static and dynamic resource allocation (e.g. via MPLS signaling protocols, ref. Fig. 2) to share architectural complexity between BB and other NEs.

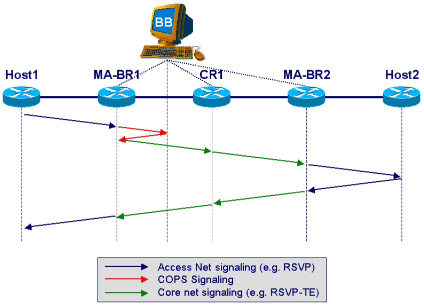

BB
Host1
MA-BR1
CR1
MA-BR2
Host2
Access Net signaling (e.g. RSVP)
COPS Signaling
Core net signaling (e.g. RSVP-TE)

Fig. 2. QoS-IP service setup in a single MAID domain.

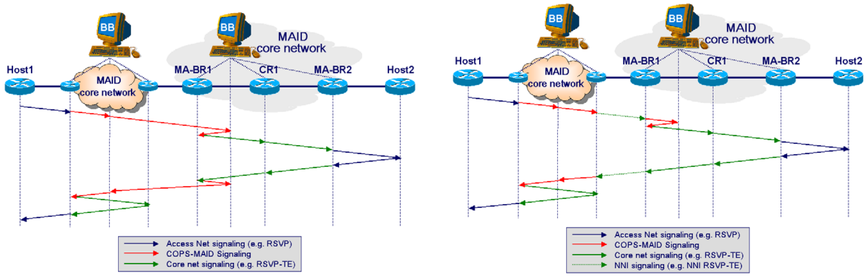
3.2 The COPS-MAID Extensions

Relying on the client-server model, the COPS architecture [10] is based on two fundamental elements: a policy server, called Policy Decision Point (PDP), also addressed as COPS server, and one or more policy clients, called Policy Enforcement Points (PEP), addressed as COPS clients. At least one policy server must exist in each administrative domain, in order to implement a complete COPS communication with one or more PEPs. A single PEP is able to support multiple client-types, while, if a client-type is not supported by the PDP, the PDP itself can redirect the PEP to an alternative PDP via COPS. Different applications using different protocols may be viewed as different client types. The trend to define a new client type for each access network protocol results in a hard limit to the system scalability, because of the duplication of the states installed both in the PDP and in the PEP. A possible solution for this issue might be the cluster of PDPs, each supporting one or few client-types; but, all of these COPS servers have either to exchange management information to perform a coherent resource allocation and should refer to a higher level ”omniscient” BB.

The novel and original solution we propose through the MAID architecture is to define a unified and extended COPS semantic, which integrate all the QoS information carried out by the different access protocols. This semantic is based on the contents of the UNI service request and it is characterized by a new unique COPS Client Type (i.e. the COPS-MAID one). The proposed extension to the standard COPS specification can be found in [11]. This solution transfers the system complexity on the border routers, in which appropriate Inter Working Units (IWUs) are used to map protocol specific messages into generalized client messages. Moreover, a unique COPS client-type can transmit all the information to a unique PDP, which can be located inside the BB itself.

3.3 The Inter-domain Problem

The research community is dealing with a number of open issues regarding inter-domain communications (e.g. the optimal TE routing, the NNI signaling, etc.). In this context, the MAID architecture arises as an effective and open solution, because of the centralized action of the BB and of the modularity of the MA-BR. Two possible strategies for the inter-domain connection setup are possible and are sketched here to prove the architectural flexibility: Inter-BB communication via COPS-MAID interface (ref. Fig. 3(a)), which has network granularity; Inter BR communications via strict NNI (ref. Fig. 3(b)), which has node granularity. The solution for the inter-domain communication among MAID domains is an inter-BB NNI, based on the COPS-MAID. As shown in Fig. 3(a), the request is processed by MA-BR similarly to the mono-domain case and, thus, propagated to the domain BB. If the destination is out of the BB scope, the COPS-MAID request is propagated to the adjacent BBs, waiting for a response. If a route exist towards the desired destination, it is announced by the downstream BB with a positive COPS-MAID response. Upon receiving this response, the BB configures the internal route from the/an ingress point towards the announced


BB
BB
MAID core network
MAID core network
Host1
MA-BR1
CR1
MA-BR2
Host2
Access Net signaling (e.g. RSVP)
COPS-MAID Signaling
Core net signaling (e.g. RSVP-TE)
NNI signaling (e.g. NNI RSVP-TE)

(a) Via COPS-MAID.

(b) Via strict NNI.

Fig. 3. Inter-domain QoS-IP service setup.

egress interface. Thus the inter-domain QoS-IP service setup is provided from the downstream towards the upstream domains. The inter-BB NNI solution proves to be more scalable than the *classical* BGP-based solution, since the space of the solution (e.g. the number of queries for a route) is limited to the adjacent BBs and not to all the possible BRs towards a different domain. A possible solution of the inter-domain communication among heterogeneous domains (such as the alternative architecture proposed in Section IV) relies on Inter BR communications. Such a solution guarantees the interoperability among architectures with centralized and distributed control planes and is still in the process of definition. A proposed approach to address this issue can be found in [12] [13].

4 The MPLS/DiffServ Backbone

In this section the key aspects of the backbone network architecture developed in the the TANGO project framework are discussed. The MPLS/Diffserv backbone here considered is able to support both pre-provisioned and on-demand LSPs establishment and to provide end-to-end LSP protection by single and double fault. Three different backbone operation models has been experimented, namely: i) centralized; ii) partially distributed; iii) fully distributed.

According to the centralized model, the control logic, needed to execute the Admission Control (AC) algorithm, the LSP route selection algorithm and the path protection mechanisms resides in a centralized device, i.e. the Bandwith Broker (BB). In this scenario the centralized device directly interacts with the backbone routers for LSP setup/tear-down and for network link status information retrieving. These interactions could be based on SNMP protocol.

In the partially distributed scheme, the control logic is still centralized, but the signaling processes for the setup and tear down of the LSPs are handled by the Edge Router (ER). In this solution, when a service request is presented at an ER, it informs the BB, via COPS protocol; the BB runs the AC algorithm and searches an available path to support the LSP request. The result of this phase are backward communicated to the ER that begins the LSP signaling phase, e.g. via RSVP-TE protocol.

In the fully distributed scheme, the control logic is completely distributed among the ERs. An ER receiving a request computes the route of the new LSP and locally performs the AC algorithm. The route selection is based on the information stored in a local Network State Database containing network topology and link available bandwidth; such information are disseminated by OSPF-TE flooding. After the route selection, the ER starts the LSP setup by triggering the RSVP-TE signalling. In this phase, AC must be performed by each node along the path because the link load information available at ER may not be synchronized with the current network state.

The following presentation is focused on the AC functions and, coherently with the scope of the paper, mainly refers to the centralized and partially distributed solutions, the specific mechanisms developed for the fully distributed solution for fault protection are given in a companion paper [14].

4.1 LSP Classes of Service

In the MPLS/Diffserv backbone defined within the TANGO project each LSP is associated to a single DiffServ PHB Scheduling Class (PSC) (corresponding to the class type defined in [15]) and to a specific protection class.

EF, AF1x and AF2x and standard Best Effort (BE) DiffServ PSC have been implemented in the test-bed, whereas, five protection classes characterized by different level of resilience and backup bandwidth sharing have been defined. Three levels of resilience have been defined: Unprotected (UP), Single-Fault Protected (SFP), Double-Fault Protected (DFP).

For UP demands only a service circuit is established, and no service continuity is guaranteed after the occurrence of a fault. For SFP demands a service circuit plus a backup one are allocated. When a failure occurs on the service circuit traffic can be readily switched on backup path. In case of DFP demands, a service circuit plus two backup paths are allocated: a primary backup and a secondary one. In this case when a failure occurs the traffic is switched on primary backup path. If a double fault occurs, it is then switched on the secondary one. Note that for DFP the preference order between the two backup paths is fixed a priori. Both SFP and DFP schemes can be implemented as Dedicated or Shared Protection [16]. In our model both Dedicated and Shared alternatives are supported for SFP and DFP, resulting in four different protection classes in addition to the basic unprotected one, as summarized in Table 1.

In the following we will index by i the DiffServ PSC, including BE. In particular, i=1,2,3,4 will refer to EF, AF1x, AF2x and BE respectively. For each

Table 1. Protection classes defined within TANGO project.

H	Acronym	Protection Class	Backup Bandwidth
0	UP	Unprotected	No backup bandwidth
1	Sh-SFP	Shared Single Fault Protection	Shared
2	De-SFP	Dedicated Single Fault Protection	Dedicated
3	Sh-DFP	Shared Double Fault Protection	Shared
4	De-DFP	Dedicated Double Fault Protection	Dedicated

generic link, we will denote by v_i the current assigned bandwidth to class i, i.e. the sum of the bandwidth values associated to all current LSPs of class i. Typically, since no bandwidth reservation is associated to BE LSPs, the assigned bandwidth is always zero for BE, nevertheless a counter has been associated to it for notation compactness. The route selection algorithm does not distinguish between the bandwidth assigned to working and backup LSPs and simply prefers the links with the largest residual bandwidth with no regards of working and backup components. Additionally we define u_i as the class i minimum guaranteed bandwidth. This is a configurable parameter that allows to enforce a minimum guaranteed cushion to class i on the specific link. Also if no class i LSPs are established, such a bandwidth cushion can not be taken by other classes and it is preserved for future class i requests. This is useful to apply bandwidth isolation between classes, which is an important requirement as stated in [17]. Whether or not apply such a minimum bandwidth cushion is a provider business policy matter. The default value for u_i is zero, except for BE class. In fact, it is likely that any provider might want to let a percentage of link capacity available to the BE traffic. From v_i and u_i, the generic node responsible for the link will extract the value of the current reserved bandwidth as

$$r_i = max(v_i, u_i) \tag{1}$$

For each LSP, the r_i counters are used to enforce local AC. The AC function must ensure that the reserved bandwidth components meet a set of constraints. These can be defined on the single values of r_i (e.g. *EF class can not exceed 50% of the link capacity*), on some combinations (e.g. *AF1x and AF2x classes jointly can not exceed 70% of the link capacity*), or on their complete sum (e.g. *the sum of reserved bandwidth for all classes can not exceed the link capacity*). Each of such constraints can be dictated by business related policies or QoS related considerations. The constraints set can be written in a formal way as follows:

$$r'L \leq c \tag{2}$$

wherein r is the column vector r={r1, r2,..} collecting the reserved bandwidth value for each class. The matrix L is composed of binary elements, each row represents a single constraint, and c is the column vector of associated limits. Similarly to r, we will denote by v and u the vectors collecting the v_i and u_i components. With the above positions, the AC algorithm can be described in a simple way. When a LSP of class j and bandwidth b is requested, the tentative value v^*_j= v_j +b (*update rule*) is computed as well as new value of vector r^*. Then, it is checked whether constraint(2) holds for the tentative vector r^*. In the affirmative case the request can be accepted and the new value of v_j recorded. Conversely, the request is refused and the counters are not updated. If backup bandwidth sharing is NOT applied, the simple update rule given above is applied to both working and backup LSPs. On the other hand, in case of bandwidth sharing, the update rule for v^*_j for the backup LSP setup must be revised according to the algorithm detailed in [18].

It is evident that, in centralized and partially distributed solutions, the centralized logic that performs the AC algorithm exactly knows the actual reserved bandwidth on each link. Vice versa, in the fully distributed solution, due to the concurrent operation of the ERs, a mechanism to continuously inform each ERs on the status of each link is needed. In such a distributed environment when a LSP is installed/removed from a link, the local node should advertise through OSPF-TE flooding the new unreserved bandwidth value for the specific link. In this way, the ERs can update their local Network State Database and perform a route selection process coherent with the current network state.

More formally, if r is the currently reserved bandwidth and c the maximum reservable bandwidth (not necessarily the link capacity), the value of g=c-r is advertised, where g represents the link residual reservable bandwidth. In order to avoid a large amount of flooding, the new value of g is not advertised upon each LSP setup/tear-down. The Opaque LSA generation process is performed according to local update policies embedding watermark-based algorithms and/or hold-down timers [19]. As a consequence, it follows the variations of g less accurately, but flooding overhead decreases. In our model we extend this approach to a multi-class environment. The *residual bandwidth* becomes a vector g = {g1, g2, ...}, whose component g_i represents the additional bandwidth that can be assigned to future class *i* requests when no new requests from other classes are performed. The computation of each component g_i involves a very simple manipulation of r, L and c. A trivial optimization problem with a single variable and linear constraints must be solved. To build the vector g, the computation is repeated independently for each class except BE. This vector can be advertised in the sub-TLV Unreserved-bandwidth of OSPF-TE Opaque LSA in conformance with the semantic defined in [20]. Given the independence between the g_i components, the same flooding reduction policy used for g in the single class environment can be straightforwardly applied to each component separately. In particular, we use the same mechanism described in [21] based on adaptive watermarks. We notice that, due to composed constraints, a LSP setup/tear-down impacts the residual bandwidth of all classes and that the sum of the components g_i can exceed the link capacity, since each element has been computed in absence of new requests from other classes. Such a semantic, coherent with the information needed by the route selection algorithm, greatly simplifies g computation and flooding reduction algorithm application. The proposed model comprises as special cases the models being currently discussed in IETF, [17] [22] [23], and is compliant with the requirements given in [22]. In particular, isolation is provided by u_i setting and inter-class sharing can be obtained by an appropriate constraint design(2). Finally, let us consider the possibility to apply TE to BE traffic. Even if no bandwidth reservation is associated to BE LSPs, it is possible to envisage a model where explicit routing capability is applied to these ones. In order to perform route selection for BE LSPs, a BE *residual bandwidth* (g_4) is defined as the *link measured residual capacity.* This value can be derived, for example, from the local MIB containing the bytes sent in the last measurement interval. In order to disseminate such an information, the value is inserted in Unreserved-bandwidth sub-TLV of OSPF-TE Opaque LSA.

5 The TANGO Platforms Assessment

In this section we describe the experimental activity aimed at assessing a functional validation of the MAID inter-domain mechanisms as well as to test data plane performance. The tests have been carried out on a distributed test-bed made up of two inter-connected domains located, respectively, in the laboratories at the Department of Information Engineering of the University of Pisa and at the META Centre of the Consorzio Pisa Ricerche. The two domains are permanently interconnected through a Gigabit Ethernet optical fiber link. At the network layer, each domain is configured as an independent autonomous system with proper strategies and policies for QoS provisioning and Traffic Engineering. The routers in each domain are prototypal routers based on IA32(PC) Linux OS platforms, equipped with the kernel modules for MPLS and Traffic Control (TC), and with the MAID-specific modules developed in the TANGO project. An overview of the test-bed topology is shown in Fig. 4. Each domain has its own BB, which manages the dynamic configuration of the network resources under its scope, as well as the inter-domain communication by means of COPS-MAID protocol. Concerning TC the scheduler used to realize the different DiffServ PSC is the Hierarchical Token Bucket (HTB) available in the Linux kernel 2.4.20. HTB is a kind of CBQ (Class Based Queuing) algorithm [24], approximating service discipline based on the class concept. The access networks/clients have been configured in order to play the role of source/destination of different kind of QoS-unaware IP traffic. Two traffic typologies have been injected in the test-bed: artificial traffic, generated by a specialized application (BRUTEv1.0 [25]), and real-time traffic, generated by the delivery of multimedia contents (based on Helix DNA platform [26]). Different tests are carried out for assessing the performance of the MAID test-bed with respect to the different source applications and traffic profiles injected into the network. These tests highlight also MAID data plane critical elements, responsible for an unexpected limitation in the overall performance. In all the tests traffic is sent after a configuration phase takes place. This phase is similar to the static resource provisioning provided by the Network Operator for those QoS-unaware access networks that can not use

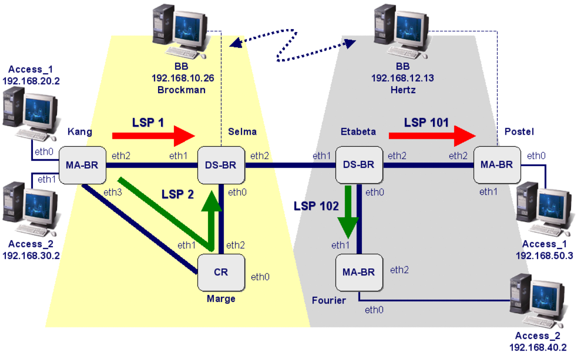


Fig. 4. MAID testbed topology.

the dynamic MAID-UNI features. Configuration consists of the DiffServ LSPs setup and traffic flows mapping into LSPs by means of a WEB interface. For each LSP, a QoS class and a reserved bandwidth are signaled. More details about these tests can be found in [27].

Artificial Traffic. The aim of this test is to identify the critical element on the MAID data plane. Two traffic flows have been generated from the same source client: the first to 192.168.50.3:6970 through an EF LSP with a reservation of 1.5Mbps, the latter to 192.168.40.2:7970 through an AF1x LSP with a reservation of 512kbps. Experiments shows that some packets are dropped even if the reserved rate equals exactly the nominal mean rate. The policer located on the ingress MA-BR is the software element responsible for this packet dropping. This element requires an accurate configuration/tuning of its parameters, in order to achieve the desired performance, above all when operating in quasi-saturation conditions. Fig. 5 shows the results obtained when two flows are generated with a Constant Bit Rate of 1.5Mbps and 512kbps, respectively.

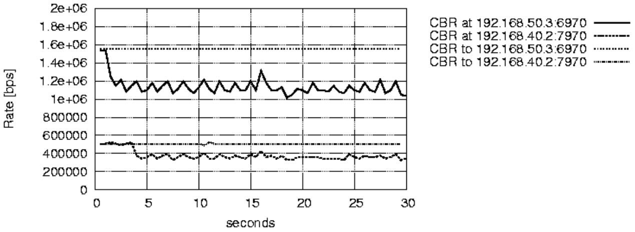
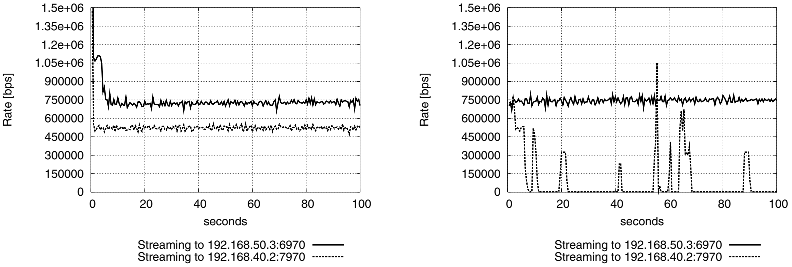


Fig. 5. BRUTE: CBR traffic flows.

Real-Time Traffic. Two types of tests are performed in this testing scenario. The first one is characterized by:

- a fixed amount of bandwidth reserved for each LSP;
- a single encoded version of the multimedia content, streamed at the encoding bit-rate of 768kbps;
- a variable connection type configured on the destination client.

The video streaming is flowed to 192.168.50.3:6970 through an EF LSP with a reservation of 1Mbps and to 192.168.40.2:7970 through an AF1x LSP with a reservation of 512kbps. Clients connection have been configured with a LAN connection speed (e.g. 10Mbps) or with a DSL one (e.g. 768 kbps). In the first case (Fig. 6(a)), after a few seconds in which some packets are dropped on both connections, due to the server attempt to fill the buffer at the full connection speed (e.g. 10Mbps), the client attached to the EF LSP perceives a good video and audio quality. Instead, the client attached to the AF1x LSP experiences a jerky reproduction because of the packet drops induced by a reserved bandwidth (e.g. 512kbps) lower than the encoding rate (e.g. 768kbps). In fact, when both clients have been configured with a DSL connection (Fig. 6(b)), no packet drops



(a) LAN speed configuration. (b) DSL speed configuration.

Fig. 6. RTSP streaming.

for the traffic flow on the EF LSP is experimented, resulting in a optimal perceived quality throughout the whole streaming. Instead, the client that receives the traffic flow on the AF1x LSP cannot succeed in filling up its buffer at an acceptable rate and it triggers automatically a PAUSE/PLAY mechanism, waiting for possibly better network conditions.

The latter type of tests is performed to evaluate performance when the same multimedia content is available at different encoding rates. In this case, the server chooses the best fitting encoding bit-rate on the basis of connection information, collected in the setup phase. These tests are characterized by:

- two different versions of the same multimedia content streamed at encoding bit-rates of 768kbps or 512kbps;
- a DSL (e.g. 768kbps) connection type configured on the clients.

In Fig. 7, the streaming is flowed to 192.168.50.3:6970 through an EF LSP with a reservation of 1Mbps and to 192.168.40.2:7970 through an AF1x LSP with a reservation of 512kbps. In this case, the clients negotiate the proper rate with the server (e.g. 768kbps for the traffic through the EF LSP and 512kbps for the other). In this situation no packet loss is experienced and the differences on the perceived playing quality are due to the different encoding rates.

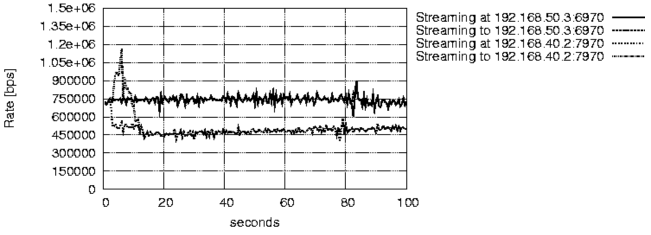


Fig. 7. RTSP streaming with different multimedia content encoding bit-rates.

6 Conclusions

In this paper a novel end-to-end seamless framework to support end-to-end Quality of Service and Traffic Engineering. The objective is to provide a complete end-to-end solution for the enforcement of a globally dynamic and multi-level chain of SLAs. Such a solution, based on the MPLS/DiffServ paradigm, allows a seamless QoS-IP service setup, according to both users and network providers requirements, by means of MAID Control Plane mechanisms. An analysis of key issues of a MPLS/DiffServ backbone have been also presented, leading to the formulation of an alternative core network architecture. The problem of intercommunication between these hetherogenous domains is still under discussion and possible solutions have been detailed. Finally, a functional validation and performance analysis of proposed architecture have been carried out on operational testbeds based on Linux PC platforms.

Acknowledgments

This work was funded by the TANGO project of the FIRB programme of the Italian Ministry for Education, University and Research.

References

1. TANGO project homepage. http://tango.isti.cnr.it.
2. S. Blake et al. An Architecture for Differentiated Services. *IETF RFC 2475*, Dec. 1998.
3. K. Poduri V. Jacobson, K. Nichols. An expedited forwarding PHB. *IETF RFC 2598*, June 1999.
4. J. Heinanen et al. Assured Forwarding PHB Group. *IETF RFC 2597*, June 1999.
5. Y. Bernet et al. A Framework for Integrated Services Operation over DiffServ Networks. *IETF RFC 2998*, Nov. 2000.
6. IST-AQUILA Project. Homepage: http://www-st.inf.tu-dresden.de/aquila/.
7. IST-TEQUILA Project. Homepage: http://www.ist-tequila.org/.
8. IST-MESCAL Project. Homepage: http://www.mescal.org/.
9. IST-MOICANE Project. Homepage: http://www.moicane.com.
10. D. Durham et al. The COPS (Common Open Policy Service) Protocol. *IETF RFC 2748*, Jan. 2000.
11. G. Sergio G. Carrozzo, N. Ciulli. COPS-MAID - COPS Usage for Multi-Access Inter-Domain MPLS-DiffServ Networks. *Internet Draft, work in progress*, Nov. 2003.
12. A. D'Achille, M. Listanti, U. Monaco, F. Ricciato, V. Sharma. Diverse Inter-Region Path Setup/Establishment.
draft-dachille-diverse-inter-region-path-setup-00.txt, July 2004.
13. F. Ricciato, U. Monaco, A. D'Achille. A novel scheme for end-to-end protection in a multi-area network. *Proc. of 2nd International Workshop on Inter-domain Performance and Simulation, IPS04, Budapest, Hungary.*, March 2004.
14. R. Albanese, D. Alì, S. Giordano, U. Monaco, F. Mustacchio. G. Procissi. Experimental Comparison of Fault Notification and LSP Recovery Mechanisms in MPLS Operational Testbeds. *submitted to QoS-IP 2005.*

15. W. Lai F. Le Faucheur. Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering. *IETF RFC 3564*, July 2003.
16. D. Awduche et al. Requirements for Traffic Engineering Over MPLS. *IETF RFC 2702*, Sept. 1999.
17. W. Lai F. Le Faucheur. Maximum Allocation Bandwidth Constraints Model for Diff-Serv-aware MPLS Traffic Engineering. *Internet Draft, work in progress*, Mar. 2004.
18. M. Listanti F. Ricciato, S. Salsano. An Architecture for Differentiated Protection against Single and Double Faults in GMPLS. *Photonic Networks Magazine*, 2004.
19. G. Apostolopoulos et al. Quality of Service Based Routing: A Performance Perspective. *SIGCOMM*, 1999.
20. D. Yeung D. Katz, K. Kompella. Traffic Engineering Extensions to OSPF Version 2. *Internet Draft, work in progress*, June 2003.
21. A. Botta et al. Traffic Engineering with OSPF-TE and RSVP-TE: Flooding Reduction Techniques and Evaluation of Processing Cost. *CoRiTeL Report*, 2003.
22. J.Ash. Max Allocation with Reservation Bandwidth Constraint Model for MPLS/Diffserv TE Performace Comparison. *Internet Draft, work in progress*, Jan. 2004.
23. J.Ash. Russian Dolls Bandwidth Constraints Model for Diff-Serv-aware MPLS Traffic Engineering. *Internet Draft, work in progress*, Mar. 2004.
24. V. Jacobson S. Floyd. Link-sharing and resource management models for packet network. *IEEE/ACM Transactions on Networking*, 1995.
25. BRUTE (Brawny and Rough UDP Traffic Engine) homepage. http://netgroup-serv.iet.unipi.it/brute/.
26. Helix DNA Server 9.0. http://www.helixcommunity.org/.
27. G. Carrozzo et al. MPLS/DiffServ interworking: preliminary functional tests for TANGO project. *TANGO project Symposium*, 2004.