# DiffServ over MPLS: Solutions and a Performance Evaluation Tool

Giulia Conte
CoRiTeL- Consorzio di Ricerca sulle Telecomunicazioni
Rome, Italy
conte@coritel.it

Paola Iovanna, Roberto Sabella
ERI- Ericsson Lab Italy
Rome, Italy
{paola.iovanna, roberto.sabella}@ericsson.com

Luigi Alcuri, Michele L. Fasciana, Francesco Saitta
DIE- Dipartimento di Ingegneria Elettrica
Università degli studi di Palermo
Palermo, Italy
{luigi.alcuri, michele.fasciana, franesco.saitta}@tti.unipa.it

***Abstract*— Next generation multi-service networks must meet basically two fundamental requirements: Quality of Service (QoS) in an Internet-based context and Traffic Engineering (TE) functionalities. DiffServ-over-MPLS stands for the most reliable solution to fulfill them in a flexible way. Anyway, even if both Diffserv and MPLS technologies have been the subjects of several works, the design and the implementation of the integrated DiffServ-over-MPLS solution, still presents unsolved nuts. This paper attempts to highlight the open issues of the integrated solution, giving an overview of the involved technologies and the possible solutions that can be implemented in such a scenario. Then a simulation tool designed to investigate such open issues is proposed. The aim is to realize a tool able to simulate different DiffServ-over-MPLS solutions so as to provide some guidelines for the analysis and the design of different network scenarios.**

***Keywords- Differentiated Services; MPLS; Quality of Service; Traffic Engineering; Network Simulator***

## I. INTRODUCTION

Next Generation Networks (NGN) will be required to fulfill two main requirements: i) the ability to support different QoS in an "all-IP" context, so as to realize real multi-service networks; and ii) the capacity to optimize the use of the network resources in order to save costs. In fact, a dilemma emerges for carriers and network operators: the cost to upgrade the infrastructure of current fixed and mobile telephone networks is too high to be supported by revenues coming from traditional Internet services. Revenues from voice-based services are currently much higher than those derived by Internet services. Therefore, to obtain cost effectiveness it is necessary to design networks that make an effective use of bandwidth or, in a broader sense, of network resources. Moreover, Internet traffic is highly variable in time compared to traditional voice traffic and is not easy to forecast. This means that networks have to be flexible enough to react appropriately to traffic changes while meet different QoS requirements for different service classes.

The employment of Differentiated Services (DiffServ) and Multi Protocol Label Switching (MPLS) techniques on the same network infrastructure seems to be the most promising solution to achieve such requirements.

DiffServ was born to provide the means to face the different traffic requirements in a scalable way for IP network. DiffServ is basically a strategic solution for differentiating packets streams that belong to class of service for which a network behavior better than best effort is requested. After having differentiated streams on the base of service class, it is possible to do an admission control specific for class of service and apply different routing strategies for each service class, in order to satisfy QoS requirements. The differentiated routing procedure can be easily implemented with the MPLS paradigm. In fact, MPLS is able to control the different traffic trunks and moreover to provide the means of achieving transmission resources optimization among the classes.

The DiffServ-over-MPLS framework has been the subject of many works in the last years: the common idea is the possibility of exploiting the benefits that come from both technologies.

In the framework of the main standardization bodies, all the key ingredients are provided to define the DiffServ mapping over MPLS, but a lot of work has to be carried out to clarify the different possibilities of inter-working between the two technologies. Actually, the QoS strategies and the TE schemes inherited from the two paradigms, give rise to several DiffServ-over-MPLS solutions. In fact, the common feature of DiffServ and MPLS is that they always offer very flexible solutions and the number of possible configurations is very high. As DiffServ and MPLS take part to the same framework, the number of tunable variables increases a lot.

The work, which this paper is a part of, aims at analyzing and preferably solving some open issues that rise from the implementation of a DiffServ-over-MPLS solution. In particular, in this paper we want to provide some guidelines in the analysis, study and implementation of different Diffserv-MPLS solutions. This is accomplished in two steps. First we give an overview about QoS and TE mechanisms provided by DiffServ and MPLS. Then we introduce a performance evaluation tool, we are developing with the goal of simulating and compare different DiffServ-over-MPLS solutions.

The paper is organized as follows.

In section II, the main features of DiffServ environment and MPLS technology are described. In particular, QoS control aspects and TE mechanisms are highlighted. In section III, the DiffServ-over-MPLS framework is presented. The aim is to show up the openings provided by the different solutions and to discuss the open issues that rise from their applications. In the subsequent section, a brief introduction of the employed tool for performance evaluation is provided. The network scenario, the protocols extensions, and the algorithm implementation are reported. In section V, we discuss the conclusions and the on-going development activity.

## II. BACKGROUND

The Internet is growing in size and structural complexity. This means that many structural changes are been studied and partially introduced into the actual framework. They mainly include:

- **Increasing connectivity:** Over the last years a steady improvement in connectivity has been observed between autonomous systems (AS) [31]. The exponential increase of last past years cause that the original backbone has turned into a backbone mesh. Introducing new management aspects that have to be investigated.
- **Basic structural changes:** More and more backbone providers enter the scene and inter-connect themselves. They also expand their geographical scope, and so causing the need to study interaction between providers management domains.
- **Increasing number of application types and protocols:** Many new applications and services require new protocols or use standard protocols in a different way than originally intended (e.g., MPLS, DiffServ). The convergence between telephony network and IP transport infrastructure sponsored by telephone operators in order to reduce costs and management has introduced the needs to offer services with different performance and Quality of Service QoS. They influence the picture when modelling traffic distributions, and new tools of analyses are requested during the project of networks

### A. *Differentiated Services and QoS control*

#### 1) *Introduction to Diffserv Domains*

A first attempt to solve the request of an IP network aware of traffic requirements is constituted by Differentiated Service, DiffServ. DiffServ is a solution for service discrimination, and locally deployable service provisioning mechanism which is gaining ground quickly. While the applications are not very sophisticated right now they are quite useful, and the introduction of DiffServ on large scale is limited by the absence of powerful tools for the configuration and managements of DiffServ into applicative network.

DiffServ based solutions, as we discuss them here, focus on the core network and assume aggregated traffic flowing to and from the access networks. A Diffserv domain is a transport subnet inside which the packet travels under the control of a specific ISP which classify them on the base of the service, which the packet belong to. In this way the ISP has the ability to separate specific streams, on the basis of destination and service level specifications (SLS), from the best-effort natural way of working of IP networks. Anyway DiffServ is a framework that has a narrow focus on mechanisms.

Only the base mechanisms and some basic forwarding types are defined as a standard [32] [33] [34] [35]. This gives a great degree of freedom to ISPs how services are provided, with which peers they interconnect, how they route traffic and what prices are applied. RFC 2475 gives a technical description of DiffServ [32]. The most important advantages of DiffServ over the other solutions are:

- DiffServ techniques scale very well, even to the size of the global Internet.
- DiffServ allows the integration of all communications between applications into one single network without splitting the available bandwidth in many small chunks, dedicated to one single application. However, if it is desired, DiffServ also allows network administrators to do the latter.
- DiffServ is easy to deploy within one Internet domain, without making the domain incompatible with other domains. It uses the existing IPv4 or IPv6 packet formats, which are currently used in almost the whole Internet.

#### 2) *Dependecies between Diffserv and QoS control*

DiffServ can be considered as a powerful tool that ISP can use in order to offer services with the benefits of QoS control and the respect of Service Level Agreements (SLA). This is accomplished by mapping SLAs into the DiffServ CodePoints (DSCP). DiffServ offers four classes of Assured Forwarding (AF) services with three different level of precedence each of them can be used for marking services that requires better performance than best-effort behavior. Moreover there is a single class of Expediting Forwarding (EF) reserved to streams for which low latency is an essential specification. A QoS control is usually applied at the ingress point of the DiffServ Domain in order to mark input packets and associate them at the appropriate flow maintaining under control the prerequisites of each different service. This is achieved with a policer applied at the edge router of any DiffServ domains that limits the number of packets marked into a same class of service. The configuration of the policer is a crucial point and with the tool described in this paper is possible to test the effects on QoS for single stream and to develop guidelines for an automatic configuration of policing mechanism.

QoS control also means an admission control of input streams as a requisite for reducing and avoiding congestion inside the transport network, the strategy adopted is a crucial issue in NGN.

Two[F.S.1] admission control strategies are proposed in literature: one based on end-to-end admission control and one based on per link bandwidth availability. In order to offer a wide possibility of configuration and the possibility of adapting the simulation tool to any applicative scenario we also propose a solution that is able to adopt and integrate both strategies inside the same framework.

*3) Application of Diffserv and QoS control in applicative scenario*

We aim to introduce a simulation tool that is able to contribute to the development and improvement of real applicative scenario, so we focus the development of this framework taking as a reference model Voice over IP as a support for telephony applications. A lot of work has been done studying the performance problems affecting the development of telephony applications over IP (ToIP) networks [36] [37].

An essential component for the success of ToIP consists in achieving the same Quality of Service (QoS) of PSTN and ISDN networks with as few changes as possible to the actual IP network implementation.

Fig. 1 shows the basic division of the network into a core part (backbones, transit networks) and access networks connecting the individual customers. The contracts between ISPs describing a peering agreement are called SLAs (Service Level Agreements). They may be based on SLSs (Service Level Specifications).

The traffic generated by the users is collected by Internet Service Provider trough their Point of Presence, which aggregate similar streams and mark them using different DiffServ CodePoints.

Each POP encloses the access interfaces for users of those services offered by Service Providers, which include real-time application such as Voice over IP (VoIP) characterized by using of different SLA.

In the case of telephony over IP, each POP is responsible for accepting new calls from the users and addressing them toward the correct POP destination granting minimum required SLSs. For this reason, in every POP there are: one or more Media Gateway Controllers (MGC), which control and allocate the resources to each call and set the connections with the other POP; Media Gateways (MG), which convert the voice communications from the PSTN/ISDN network into IP packets and collect statistics for each call received; and Internet Gateway with the functions of monitoring and marking the IP packets received by users on a local LAN, which already adopts Voice-IP devices inside their organization.

One of the main advantages of using ToIP is the possibility of implementing different telephony services, which have different costs in relation with the effective network resources requirement and utilization. Each SP has to characterize its services with well-defined Service Level Agreements (SLA).

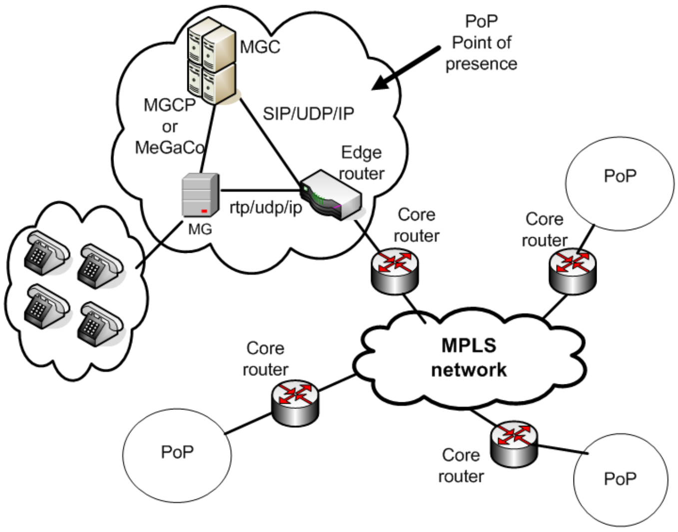


Figure 1. Network scenario for TOIP QoS controlled phone calls

We chose to implement several services with different specifications; each service is focused on a specific feature that could be considered relevant for the development of future ToIP applications.

- Voice over IP services are enclosed inside DiffServ Assured Forwarding classes and are mainly focused on guaranteeing latency requirements for packet streams and packet loss rates.
- File transfer services, which could be used for some non real-time telephony applications such as fax, sms, e-mail or file transfer, are focus on the possibility of give to this services a minimum guaranteed bandwidth. The Data service is intended to study the effects of aggressive UDP privileged streams versus connection-oriented stream with auto-control congestion avoidance mechanisms, such as TCP connections.

*B. MPLS and TE solutions*

In the panorama of technology solutions provided to face the aforementioned NGN requirements a key role is played by MPLS and TE. They basically offer the capability of dynamically routing the traffic over the network in order to minimize congestions and to optimise the use of network resources, while at the same time guaranteeing a certain grade of service, handling traffic fluctuations and offering multi-service capabilities ) [10] [11].

In this section a brief introduction to the MPLS technique and its TE applications is given.

*1) MPLS*

MPLS architecture is a standardized structure based on the separation between data plane and control plane, that reuses and extend existing IP protocols for signaling and routing functions, while reintroduces a connection-oriented model in an Internet-based context [12]. The MPLS scheme is based on the encapsulation of IP packets into labeled packets that are forwarded in a MPLS domain along a virtual connection named Label Switch Path (LSP). MPLS routers are named as Label Switch Router (LSRs) and the LSRs at the ingress and at the egress of an MPLS domain, are named Edge-LSRs (E-LSRs). Each LSP can be set up at the ingress LSR by means of an ordered control, before packets forwarding. That LSP can be forced to follow a route that is calculated a priori thanks to the explicit routing function. MPLS allows the possibility to reserve network resources on a specific path by means of suitable signaling protocols (e.g. RSVP-TE [13], CR-LDP [14]). Thus, the LSP represents a virtual connection in the MPLS network as the virtual circuits and virtual paths are in the ATM world.

In particular, each LSP can be set up, torn down, re-routed if needed, and modified by means of the variation of some of its attributes, including the bandwidth [12]. Furthermore, pre-emption mechanisms on LSPs can also be used in order to favor higher priority data flows at the expenses of lower priority ones, while avoiding congestion in the network. For that reason, each LSP is associated to a level of priority that determines the degree of precedence in resource reservation.

Another important feature of MPLS relates to the possibility of stacking labels that provides the means to introduce different hierarchical levels instead of the two ones provided by ATM [11]. This feature favors VPN services support.

The strength of TE within MPLS is the ability to control virtual connections (i.e. LSP) dynamically by means of IP based protocols. This allows introducing a high level of flexibility that is proper of IP world.

In order to better explain how TE within MPLS works, key MPLS functions, such as Constraint Based Routing (CBR), resource reservation mechanisms and admission control, and their inter-working are described in the following.

*2) MPLS key functions for TE and related protocols*

One of the key functions that MPLS uses for TE is the constraint based routing (CBR) [11]. CBR is the capability to perform route calculation taking into account different constraints instead of the number of hops as in the case of plain IP routing. In particular, the criteria utilized to choose routes in a network and, possibly, to re-route traffic flows towards alternative paths, are crucial for applying TE strategies. Such criteria necessarily take into account more parameters than simply network topology. Specifically, when calculating the route for a LSP, CBR has to take into consideration both network and user constraints. The former regards the link state, resource availability besides network topology, while the latter relates to bandwidth requirements, administrative groups, priority etc. This procedure allows to perform more intelligent routing with respect traditional shortest path routing (e.g. traditional OSPF). As a result CBR can consider longer but less congested paths instead of heavily loaded shortest paths. Thus, network traffic is distributed more uniformly and congestions are prevented.

The other key capability of MPLS is the possibility to perform the bandwidth reservation by means signaling protocols such as LDP and RSVP that are suitable extended to support such functions. In order to evaluate the inter-working between routing and signaling protocols, let us consider the control plane of a MPLS network with Traffic Engineering functional. Each node supports both a routing protocol and a label distribution protocol. The possible routing protocols are OSPF-TE [22] and ISIS-TE [15], which extend OSPF and IS-IS respectively. Specifically, the traditional routing protocols have been enhanced with the ability to carry information related to link attributes/states, to be used for explicit route calculation (e.g. available/reserved bandwidth). The label distribution protocol (or "signaling" protocol) is used to setup the LSPs, supporting both explicit route indication and reservation of resources during dynamic LSP setup. RSVP-TE and CR-LDP are the two "TE-capable" label distribution protocols. The setup (and release) of LSPs modifies the resource allocation status on the network links. Therefore the signaling protocol interacts, for instance, with OSPF-TE within a node to communicate the resource allocation status. Once OSPF-TE has been notified of a change in the status, it should advertise this change to all other LSRs by sending a Link State Update (LSU) message containing a special kind of Link State Advertisement (LSA) object called opaque LSA. The object is called opaque because it is "hidden" to the basic OSPF routing logic, as it is only used by the TE logic.

After route computation, MPLS provides LSP establishing by RSVP that is suitable extended to support the information about the route that has to be established.

The path reservation is performed link by link and a local admission control is achieved to check the availability of bandwidth on the specific link in order to accommodate the LSP requested. During the local admission control, besides the bandwidth availability, also the LSP priority attribute is taken into account. Clearly this step, that is proper of oriented connection architecture, is a key mean for MPLS to provide service guarantee.

It is worth to highlight that the admission control function in a MPLS contest can take a different mean with respect to the role that assume in case of DiffServ environment, where the resources assignment is performed at the edge of the network or in centralized unit [16]. The inter-working between admission control in DiffServ and MPLS is an open issue better described in section III C.

## III. DIFFSERV-OVER-MPLS

### *A. MPLS support of DiffServ*

As mentioned before, in multi-services IP networks, whereas QoS control and resources optimization is required, Diffserv architectures may be employed over MPLS TE mechanisms. In such a framework TE mechanisms are thought to operate on an aggregate basis across all DiffServ classes of services.

In [2] a solution for DiffServ support in MPLS networks is provided. Basically, the solution gives the means to choose the mapping between DiffServ classes and MPLS LSPs, in order to perform TE strategies according to traffic QoS requirements. Two kinds of mapping are defined:

- E-LSP (EXP-inferred-PSC LSP): LSP able to carry up to 8 DiffServ Behavior Aggregates, whose PHBs are specified in the EXP field of the MPLS header;
- L-LSP (Label-only-inferred-PSC LSP): LSP able to carry a single DiffServ Ordered Aggregate, whose PSC is specified by the label. The EXP filed refers only to the drop precedence.

More intuitively, L-LSP supports a single class of service, allows traffic engineering at a very fine-grained level and simplifies VPN creation/management: routing, protection, re-routing, and pre-emption are separated for each service class. In contrast, E-LSP is able to support a set of service classes; they are traffic engineered (routed, protected, restored, etc.) all together, thus providing less flexibility. However E-LSPs could perform better in terms of scalability; in fact, the management of a single E-LSP encompassing N classes (instead of N L-LSPs) requires less signaling and smaller routing tables.

The mapping between the MPLS label and the correspondent DiffServ class(es) can be explicitly signaled during the LSP set-up. In this case, MPLS signaling protocols, LDP and RSVP-TE, must employ the Diffserv TLV and DIFFSERV object, respectively, defined in [2].

As in the existing MPLS TE, the LSP set-up can be performed with or without bandwidth reservation. In the first case the signaled bandwidth may be used by the LSRs to execute admission control during the path set-up over the Diff-Serv resources provisioned for the correspondent PSCs. Moreover, there is the possibility of using such signaled bandwidth to adjust the resources allocated to the relevant PSCs before performing admission control.

If admission control in MPLS becomes dependent on bandwidth segregation among different classes, there is an implicit impact in path selection algorithms.

QoS routing has already introduced the DiffServ perspective in path selection, by finding routes that are most likely to be able to meet QoS requirements [9]. For instance, different objective functions can be employed for different classes of service [28] (e.g. minimize hop count for delay-sensitive traffic trunk). This is the first effort to achieve the so-called “end-to-end QoS”.

The Constraint-based Routing, the most powerful tool employed by MPLS to execute TE, evolves QoS Routing by adding further network constraints in path calculation (i.e. network link attributes). Thus, in the DiffServ-over-MPLS scenario, besides the different objective functions, different bandwidth constraints according to different classes of service may be used in path selection.

### *B. DiffServ-aware MPLS Traffic Engineering*

Performing per-class TE, rather than on per-aggregate basis across all classes, is the foundation which the Diffserv-aware MPLS TE (DS-TE) [3] is built on.

DS-TE is based on the possibility of splitting the traffic according to the classes of service, into multiple traffic trunks which are transported over separate LSPs. The LSP path selection algorithm and the MPLS admission control procedure can take into account the specific requirements of the traffic trunk transported on each LSP (e.g., bandwidth requirement, preemption priority). Such requirements are translated in engineering constraints applied to the routed traffic trunks.

DS-TE solution basically has two goals: i) limit traffic trunk transporting a particular class to a relative percentage on core links, ii) more efficiently use network capacity by exploiting MPLS TE peculiar functionalities (e.g. preemption).

Some examples of specific environments, which would benefit from DS-TE are cited in [3] and embrace networks with limited bandwidth capacity (e.g., transcontinental networks), networks carrying a significant amount of delay-sensitive traffic, and networks where the entering traffic is not uniformly distributed across service classes.

For link bandwidth allocation, constraint based routing and admission control, DS-TE makes use of the concept of Class Types (CTs) previously defined in [1]. A CT is a set of traffic trunk crossing a link ruled by a specific set of bandwidth constraints. DS-TE must support up to 8 CTs. The Diffserv classes can be mapped to CTs without any particular limitation. For instance, it is possible to associate two CTs to the AF1x and the AF2x PSCs, as well as to consider the two traffic trunks in a single CT taking into account the similar nature of their QoS objectives [3].

The set of bandwidth constraints and how they govern the CTs are defined in the Bandwidth Constraint (BC) Model. Some models have been defined: Russian Dolls Model [5], Maximum Allocation Model [6], Maximum Allocation with Reservation Model [7]. The aim of the BC models is, roughly speaking, to regulate the amount of traffic related to different CTs across network links. Some models present a higher sharing degree of link bandwidth among the different classes while others provide stricter class isolation. Of course the best solution depends on the network condition [8] (overload/normal condition, relative proportion of class of services entering the network…).

Since preemption priority associated to an LSP results independent from CT definition, it is possible to characterize an LSP by its preemption level and the CT it belongs to. In DS-TE framework, a traffic class identified by a given CT and with a certain preemption priority is called “TE-Class”.

In order to support such new TE-Classes, extensions to MPLS routing and signaling protocols have been already proposed [4]. Regarding RSVP-TE, a new optional object, named CLASSTYPE object, has been introduced in the Path message. During the path establishment the CT attribute is used in the admission control procedure in addition to the preemption attribute. Regarding IGP protocols, new optional sub-TLV has been defined in order to carry the Bandwidth Constraint Model and the Bandwidth Constraints applied to the advertised link. The existing “Unreserved Bandwidth” sub-TLV, defined in [24], is used to carry the information about the unreserved bandwidth for each TE-Class. Such information is used in the constraint based route calculation. In general, the Unreserved Bandwidth can be a function of the actual reserved bandwidth of each CTs, the BC model, the BCs, and the maximum reservable link bandwidth (that can be more/less the actual link capacity in case of overbooking/underbooking). In [5] [6] example formulas for computing Unreserved Bandwidth for a given TE-Class are shown. Anyway a fundamental requirement is that the unreserved bandwidth values must be consistent with the Bandwidth Constraint model. Roughly speaking, the admission control algorithm and the path selection procedure have to follow the same criteria.

### *C. Open Issues*

The different solutions prospected in the DiffServ-over-MPLS framework present several interesting open issues.

For instance, regarding the DS-TE context, [8] shows how bandwidth constraint models present the trade-off between bandwidth sharing, to achieve better efficiency under normal condition, and class isolation under overload network condition. Thus, besides the investigation of new models, it is possible to study such a trade-off employing different routing and admission control algorithms. In fact, within the aforementioned models, methods for calculating unreserved bandwidth are proposed, but they are only examples. The only requirement is that the unreserved bandwidth values (employed in path selection) and the admission control procedure must reflect the same bandwidth constraint model.

Moreover, in all the Diffserv-over-MPLS solutions, there is the optional possibility to dynamic adjust Diffserv PHBs parameters (e.g. queue bandwidth) during the establishment of a new LSP. This option may involve a number of configurable parameters: an investigation about the trade-off between complexity and performance gain could be the matter of a research study.

Another open issue is how to map PHBs in CTs, and how to associate preemption priorities to form TE-Classes in the DS-TE framework. The requirement is that a solution is able to configure up to 8 TE-Classes. Anyway, since the number of CTs can be up to 8 and the preemption priorities go from 0 to 7, there are at disposal 64 possible combinations. When choosing a sub-set of these combinations to configure LSRs, the interaction of classes at Diffserv level has to be considered when setting the preemption priorities. In fact, there are TE mechanisms, such as soft-preemption [17], where reservation contention may not reflect forwarding plane congestion.

Another point of the integration that raises some concerns is the admission control procedures. Both the technologies have their rules that they use for accepting or rejecting traffic. Are they complementary or exclusive? If LSPs are established with resources reservation, an admission control node-by-node along the path is performed: is DiffServ edge admission control still necessary? It will become integrated with the MPLS procedure? It will trigger LSP set-up becoming the intelligence of the network?

The investigation of the presented open issues (or a part of them) is enough complex since it requires the analysis of aspects regarding both the control and the forwarding plane. Furthermore, the number of variable that can be tuned could increase as more as the DiffServ-over-MPLS solution is sophisticate.

A tool that is able to simulate from the simplest TE scenario to the most complex DS-TE solution would be very useful for this purpose. Moreover, the prospective to provide trends to design VPNs in a Diffserv-MPLS environment makes the employment of a simulating tool very attractive.

Network Simulator, NS2 [18], is an open-source simulator that fulfills such research requirements.

Besides the possibility of extending the already implemented protocols to actualize TE functionalities, NS results very useful to study the different behavior of the DS-TE solution under different boundary condition. With NS it is also possible to model different traffic sources from simple real-time voice applications to more sophisticated video conference and video streaming applications; that belong with the possibility of simulating different clients at the edges of the network makes this tools useful for the study of a large number of different network scenarios.

## IV. A Tool for Investigation and Pervormance Evaluation[F.S.2]

### *A. DiffServ module and its interaction with Application and MPLS layers*

#### *1) From the definition of Service Level Agreements to Diffserv Codepoints and Per Hop Behaviours*

As a first step in the definition of simulation scenario it is crucial to individuate the services to offer at the user and the corresponding service level specification. This tool makes use of NS2 DiffServ module in order to differentiate the input streams on the base of the service, which belong to, and matching to relative service level specifications.

After having established the kind of service which have to be managed inside the DiffServ domain it is necessary to assign them the corresponding DSCP, which will mark all the packets which belong to a specific service and influence the Per Hop Behavior (PHB) of DiffServ capable routers.

We can differentiate the service on the basis of their performance requests described by SLA and SLS in terms of latency, packet loss rate and bandwidth.

For each class of service individuated there is an assignment inside a forwarding class of diffserv, such as presented in the following table.

Our framework is configured for adopting Expediting Forwarding for the service of control packets such as SIP signaling and MPLS routing and signaling messages; while Best Effort for data traffic such as File Transfer application.

AF classes are utilized by real-time application, for example we have configured different voice rate services each with different voice compression codec.

The DiffServ module is used on each router at the edge of Diffserv Domain for marking ingress packets to each corresponding PHB by means of Diffserv CP. The marking is effectuated at each edge router by Token Bucket policers and all the traffic in excess goes to a low priority precedence class, as described in table 1.

TABLE I ASSOCIATIONS OF SERVICES TO DIFFSERV CLASS

| DiffServ Class | Drop Precedence level | | |
|---|---|---|---|
| | Gold (1) | Silver (2) | Bronze (3) |
| EF | UDP Control Packet | | |
| AF1 | Voice (64k) | … | Voice overload |
| AF2 | Voice (32k) | … | Voice overload |
| AF3 | Voice (8k) | … | Voice overload |
| AF4 | Voice (5k) | … | Voice overload |
| Best Effort | Data traffic | | |

Each forwarding class is managed on the basis of precedence level associated to router physical and virtual queue for simulating the difference precedence level implemented in DiffServ definitions.

*2) Diffserv Edge router Queue Scheduling*

Queues are managed using a Random Early Drop (RED) algorithm, in order to reduce the mean queue length and to avoid burst losses during network overloads.

At this point we introduce the queue scheduler which has been configured to manage two physical queues: one for UDP voice packets and one for Best Effort data traffic, in order to manage different queue priorities and minimum throughput requirements for services level agreements. A special queue with maximum priority is reserved to EF class.

Different algorithms can be tested inside the framework in order to evaluate and improve the performance of the network, actually the following algorithms are implemented:

- **(RR) Round Robin**: this algorithm is the most simple for the management of multiple queues, and represent a bound on the advantages of queue scheduling into the scenario considered.
- **(WRR) Weighted Round Robin:** This algorithm gives a weight to the different queues in order to give proportional service time at each queue.
- **(WIRR) Weighted Interleaved Round Robin:** This algorithm works like WRR but introduce an Interleaved behaviour, which permit the service of queues out of their service time, when there are no packet to serve in the actual serving queue.
- **(PRI) Priority Queue:** In this case, it is possible to configure a service priority to the different queues according to their specific kind of service
- **(SFQ) Stochastic Fairness Queuing:** This algorithm works like round robin introducing a stochastic perturbation in the service of the queues maintaining the fairness between the queues

*3) Diffserv Edge routers and End-to-end call admission control*

Ingress edge router operates a preliminary end-to-end admission control evaluating the transport performance of each real-time stream inside the DiffServ Domain. As a prerequisite for effectuating such kind of control we have developed a new NS module which is able to collect statistics of real time streams and predict the status of congestion of transport network on the basis of preconfigured thresholds. At this stage of development is possible to collect statistics about number of packets transmitted, number of packets received correctly, number of packet loss with related packet loss rate and maximum burst loss, minimum mean and maximum propagation time belong with packet arrival jitter. This module simulate the functionality of MGC of collecting statistics from MG interface by means of MGCP or MeGaCo protocol in order to monitor the status of connection and to act behaviour improvement. The knowledge of this parameters is the base for the introduction of several Call Admission Control mechanism. Even though at the present stage of development is possible to signal a network congestion on the basis of number of calls and bandwidth availability at the edge routers, in the next step of development we aim to introduce more sophisticated Admission control algorithims, such as:

**Simple Measured**: an admission control algorithm, that simply ensures that the sum of requested resources does not exceed link capacity.

**Measured Sum**: an admission control algorithm that uses measurement to estimate the load of existing traffic.

**Equivalent Bandwidth**: the equivalent bandwidth [38] of a set of flows is defined as the bandwidth such that the stationary bandwidth requirement of the set of flows exceeds this value with probability at most $\alpha$.

**QoS-Weighted Bandwidth**: a CAC algorithm in which the call admission was managed by a self adaptive control mechanism [39], which pointed to reach the following aims:

- To guarantee that every service was not affected by the traffic generated by other service.
- To guarantee that the final QoS of voice calls was better than a minimum level, which is determined by the following parameters: latency, jitter, packet loss rate and maximum burst loss.
- To guarantee that the throughput of the bottleneck was very close to the theoretical maximum of the link.

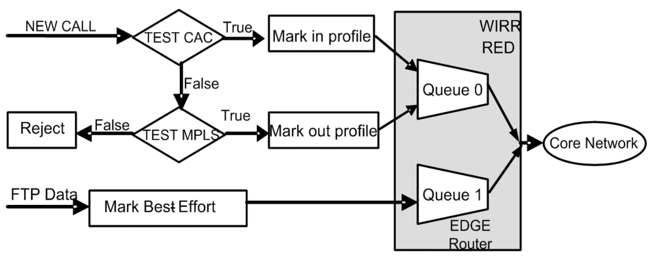


Figure 2. Block-diagram, which illustrates the admitting and marking of new calls.

#### 4) Diffserv Interaction with MPLS and Reserved Path

The DiffServ-MPLS integrated scenario present an interesting solution to control and manage network congestion related to the possibility of acting on packet routing before of blocking new incoming calls. In this case the signals of network congestion adopted for refusing call at the edge router are adopted to trigger the core MPLS router in order to establish a new LSP with better characteristics able to serve the class respecting the foreseen service level specifications.

In more details (figure 2), a new call request starts a check of call admission mechanism that analyze the current status of transport network by means of the statistics collected by MG interfaces for the specific service class of the new call. Before blocking a new call on receiving error message of insufficient service, the MGC query to ingress MPLS router for instauration of a new LSP for the aggregate flow of the service class including the new call request. In the case of successful response the new call is admitted and the overall streams is routed trough a new path with more resources available.

In this way, the combination of Diffserv CAC and MPLS-TE from an user point view improve the quality of service perceived and from a ISP point view helps to maintain balanced the load throughout the network improvinc efficiency and performance.

### B. Implementation of MPLS TE extensions

The implementation of MPLS TE functionalities in the simulator has been dived in two fundamental steps. The first regards the development of the basic TE functionalities: LSP set-up, bandwidth reservation, constraint-based routing. The second step considers all the functionalities that regard the interoperability with the DiffServ implementation part, as described within the DS-TE context. At the moment, we are able to present the implementation regarding the first step.

A module implementing MPLS functionalities is available embedded in the NS2 distribution. It is derived from a contribution code, called MNS (MPLS Network Simulator) [19], and provides a certain set of TE functionalities. Herein the employed signaled protocol is LDP. From such implementation, we reutilize the label switching functionalities for the forwarding plane and the label distribution procedure for the control. The functions of admission control and resource reservation have been implemented separately. We have chosen not to exploit the already implemented procedures because we mean to implement different strategies (different PHB-to-CT mapping, different algorithms, different booking schemes). Now the procedure of admission control has only one class type that represents the aggregation of the four real-time voices classes entering the network.

Regarding the routing protocol, TE functionalities have been added to the OSPF implementation. The starting module is the QOSPF module for NS available in [20] and implementing functionalities describe in [23]. We have extended this module in order to simulate the TE LSA (specified in [24]) and to add an our constraint-based routing algorithm. TE LSAs are Opaque LSA carrying information about the traffic engineering topology (including bandwidth and administrative constraints). The implemented constraint-based routing has the objective of balancing as well as possible the load in the network. Since, as aforementioned, the employed admission control only support one class type and since the routing algorithm and the admission control procedure must reflect the same bandwidth constraints, the implemented algorithm is based on the aggregate bandwidth associate to the involved traffic. The result is that the percentage of Assured Forwarding classes across the core links is maintained under control.

Since the next step will be the support of multiple CTs, it will be possible to create bandwidth constraints also for data traffic. This will lead to a scenario where all the traffic flows enter the network are governed by engineered constraints. Moreover, the four Assured Forwarding classes may be mapped into different TE-Classes, and thus they may get different services from the connection priority point of view.

### C. A client scenario: SIP signalling among MGC

The possibility to accept or to refuse access request, through mechanisms of CAC and to differentiate the packets related to different connections, through marking, are characteristics that have to be in the Next Generation Networks.

In this scenario decomposed gateway model is inserted: a Media Gateway Controller (MGC) [45] in which the intelligence of the whole system resides, and that, through Media Gateway Control Protocol (MGCP), manages an Access Gateway (AG) and receives legacy call signalling (QSIG over IP) from a Signalling Gateway (SG). Local MGC dialogues with remote MGC through Session Initiation Protocol (SIP) [40].

Through MGCP commands, MGC can order to AG of monitoring any connection it likes. Furthermore MGC can effect the most opportune choices for everyone that requests an access to IP network. Particularly, in case of VoIP applications albeit is possible, through MGCP, to:

1. monitor the state of the network consulting dedicated databases, continually updated through the information encapsulated in MGCP messages;
2. activate QoS mechanisms (marker and policy) on a connection, based to management policy or actual conditions on the network;
3. reject access requests;
4. monitor the activated connections.

*Call Signalling in Next Generation Telephone Service*

SIP is an application-layer control protocol that can establish, modify and terminate multimedia sessions or calls. These multimedia sessions include services like multimedia conferences, Internet telephony and similar applications. SIP is one of the key protocols used to implement VoIP.

QSIG is a signalling protocol that operates between Private Integrated Services exchanges (PINX) within a Private Integrated Services Network (PISN). In the last times many PISN adopt IP-based backbones to forwarding own data and voice traffic (Virtual Private Networks). But, while it is possible to use standardize techniques for forwarding voice payload (voice samples), there is no even a reliable standard for signalling transfer and this is an open issue.

SIP-T [41] refers to a set of mechanisms for interfacing traditional telephone signalling with SIP. The purpose of SIP-T is to provide protocol translation and feature transparency across points of ISUP-SIP or QSIG-SIP interconnection. The SIP-T effort provides a framework for the integration of legacy telephony signalling into SIP messages. To reach the goal, SIP-T provides techniques known as *encapsulation and translation* respectively [43].

Figure 4 shows a network configuration using SIP-T. In this example Access Gateways (AGs) are connected to the Private Switched Circuit Network (PSCN) via E1 carriers. Media Gateway Controllers (MGSs) manage AG by MGCP or MeGaCo Protocol. The Ingress Media Gateway Controller (I-MGC) may receive a signaling call from Signalling Gateway (SG) connected to PBX. The signaling information from the carrier must be processed by I-MGC to establish the originating call half, and to determine the identity of the Egress Media Gateway Controller (E-MGC) required to complete the call. I-MGC uses SIP to communicate the necessary information to the E-MGC to complete the call. E-MGC is able to establish the terminating call half on any of the supported trunk types.

At the edge of the depicted network, an MGC converts the QSIG signals to SIP requests, and sends them on other MGCs. Although figure 4 depicts only two MGCs, VoIP deployments have many such points of interconnection with the PSTN and VPN.

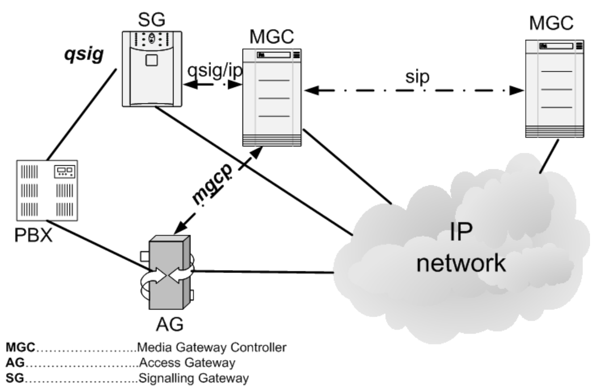


Figure 3. Decomposed Gateway

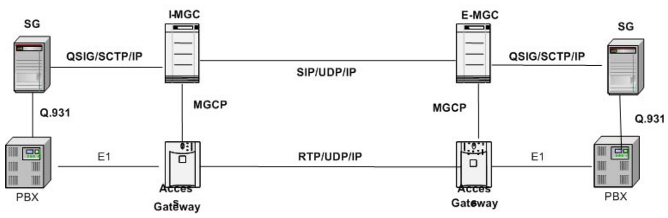


Figure 4. Decomposed Gateway: Protocol

Figures 5 and 6 show two call flows related SIP-T with voice calls originate and terminate in the VPN (via gateways). In the first figure, we show the case of a call setup with regular access (good end of CAC procedure); in the second, remote MGC sends an SIP error message because problems occur in the network.

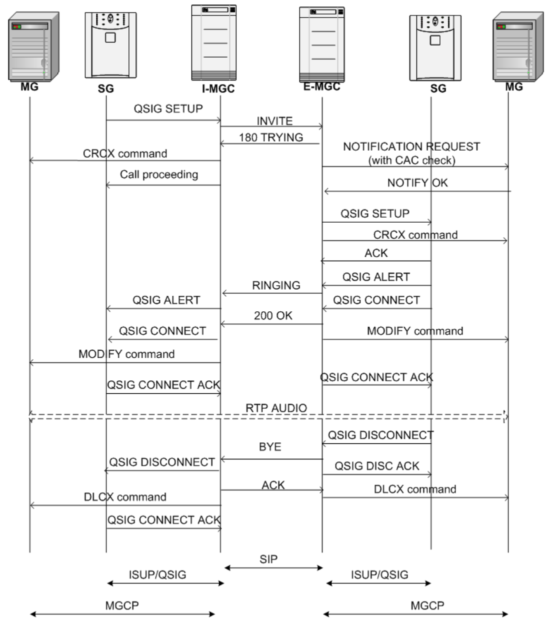


Figure 5. SIP-T call flow: normal set-up

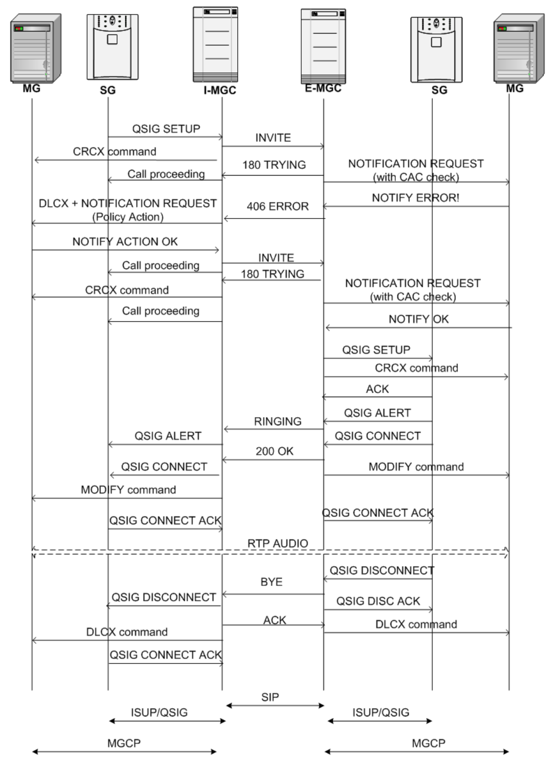


Figure 6. SIP-T call flow: set-up with QoS network failure

### *MGC-module for NS2*

We have projected a new tool for NS2, able to simulate MGC behavior in environments like those described in fig. 7.

MGC module provides to simulate a bidirectional dialogue between two SIP agents that send INVITE requests. Only when call setup is correctly performed, a new bidirectional voice session starts between two nodes. Then developed voice traffic depends on the number of call setup correctly performed.

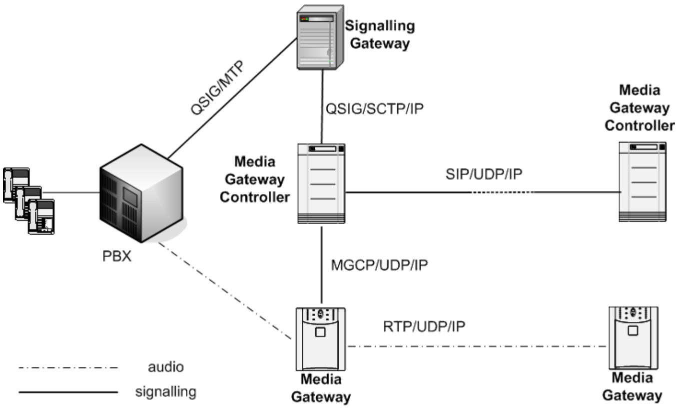

Figure 7. MGC module: reference scenario

### *MGC-MGC dialogue : SIP*

The inter-arrival process between two consecutive sent INVITE messages is poissonian. According to call flow described in figure 8, the TRYING message is sent instantly as soon as an INVITE message is received; timers manage the forwarding of SIP messages properly: RINGING timer (provides uniform delay) and 200OK timer (provides exponential delay). Finally BYE timer controls the stop of the single call (the time is stochastic too).

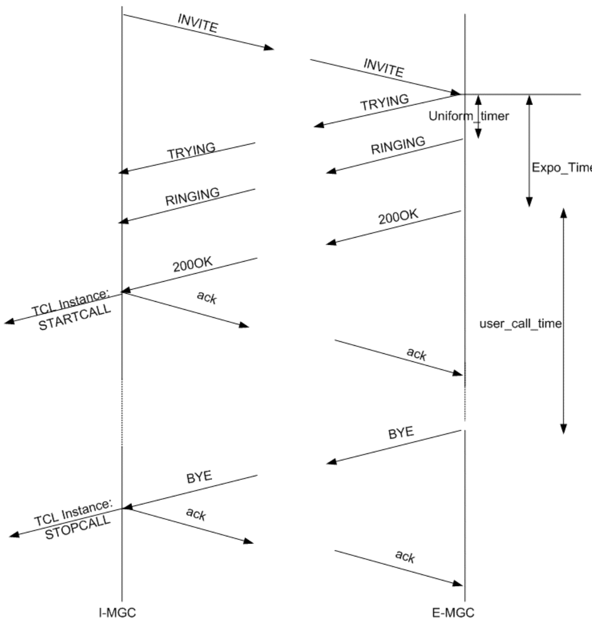

Figure 8. MGC Module: timers

### *MGC-MG dialogue :MGCP*

Between one SIP message and the following one, MGC sends a series of command to MG using MGCP protocol. Among the problems that the simulation of the scenery SIP-T introduces, there is that to synchronize the departure of the vocal calls (in the reality manages through MGCP) with the instant in which MGC agents receives the message 200OK; in fact, it is just in this moment that all the fit procedures to the installation of the call (the information exchange through SDP, opening of the doors RTP and UDP in the Media Gateway, choice of the codecs, etc.) are dispatched.

NS2 provides the possibility to use particular procedures able to modify the carrying out of the simulation (execution of the script tcl) directly from the rising C++ code. Several Tcl procedures (Insproc) are been inserted for managing voice sources during the simulations. In particular:

- *Instproc configure session:*
  it refers to CRCX MGCP command
- *Instproc OK CAC:*
  it refers to NTRQ MGCP command (check resource)
- *Instproc check MPLS:*
  it refers to NTRQ MGCP command
  (to query MPLS for a new LSP)
- *Instproc start call:*
  it refers to MDCX MGCP command (starting call)
- *Instproc stop call:*
  it refers to DLCX MGCP command (end call)

In this way MGCs can manage autonomously set up, start and stop of the source voices.

## V. CONCLUSION

This paper reviews the DiffServ and MPLS models, and discusses the architectural framework of DiffServ-over-MPLS, highlighting the open issues that still have to be addressed. In particular, a performance evaluation tool, based on simulation over the NS platform, is proposed. The aim of such a tool is the analysis of the main solutions that address the formerly individuated open issues such as the admission control and bandwidth constraint models in the integrated system. The proposed simulation tool allows all the different QoS aspects, either relating to the control or the data plane, to be taken into account simultaneously in the same simulation context. Moreover the modular structure of NS provides the means for adding new modules to implement new functionalities.

## REFERENCES

[1] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, and X. Xiao, “Overview and Principles of Internet Traffic Engineering,” RFC 3272, May 2002

[2] F. Le Faucheur, et al., “Multi-Protocol Label Switching (MPLS) Support of Differentiated Services,” RFC 3270, May 2002

[3] F. Le Faucheur, W. Lai, “Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering,” RFC 3564, July 2003

[4] F. Le Faucheur, “Protocol extensions for support of Diff-Serv-aware MPLS Traffic Engineering,” draft-ietf-tewg-diff-te-proto-05.txt, IETF Internet Draft, work in progress, September 2003

[5] F. Le Faucheur and W. Lai, “Maximum Allocation Bandwidth Constraints Model for Diff-Serv-aware MPLS Traffic Engineering,” draft-ietf-tewg-diff-te-mam-01.txt, IETF Internet Draft, September 2003

[6] F. Le Faucheur, "Russian Dolls Bandwidth Constraints Model for Diff-Serv-aware MPLS Traffic Engineering," draft-ietf-tewg-diff-te-russian-04.txt, IETF Internet Draft, September 2003

[7] J. Ash, "Max Allocation with Reservation Bandwidth Constraint Model for MPLS/DiffServ TE & Performance Comparisons," draft-ietf-tewg-diff-te-mar-02.txt, October 2003

[8] W. S. Lai, "Bandwidth Constraints Models for Differentiated Services-aware MPLS Traffic Engineering: Performance Evaluation," draft-wlai-tewg-bcmodel-03.txt, IETF Internet Draft, September 2003

[9] Xiao, Ni, "Internet QoS: a Big Picture," IEEE Network, March/April 1999, pp. 8-18

[10] X. Xiao, et al. , "Traffic Engineering with MPLS in the Internet", IEEE Network, March/April 2000, pp. 28-33.

[11] D. Awduche et al., "Requirements for Traffic Engineering over MPLS", IETF RFC 2702, September 1999.

[12] E. Rosen, et al. " Multiprotocol Label Switching Architecture", IETF, RFC 3031, January 2001

[13] D. Awduche et al, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[14] B. Jamoussi, et. al., "Constraint-Based LSP Setup using LDP," RFC 3212, January 2002

[15] A. Zinin "Cooperative Agreement Between the ISOC/IETF and ISO/IEC Joint Technical Committee 1/Sub Committee 6 (JTC1/SC6) on IS-IS Routing Protocol Development", IETF RFC 3563, July 2003

[16] P. Trimintzios, et al. "Engineering the Multi-Service Internet: MPLS and IP-Based Techniques" Proceedings of IEEE International Conference on Telecommunications, Bucharest, Romania, June, 2001.

[17] M. R. Meyer , Denver Maddux , J. P. Vasseur, C. Villamizar, and A. Birjandi, "MPLS Traffic Engineering Soft preemption," draft-ietf-mpls-soft-preemption-01.txt, IETF Internet Draft , October 2003

[18] "The Network Simulator - ns-2", http://www.isi.edu/nsnam/ns/

[19] "MPLS Network Simulator", http://flower.ce.cnu.ac.kr/~fog1/mns/

[20] "QoSR ns-2" http://www.tct.hut.fi/tutkimus/ironet/ns2/ns2.html

[21] J. Moy, "OSPF Version 2," RFC 1583, March 1994

[22] R. Coltun, "The OSPF Opaque LSA Option," RFC 2370, July 1998

[23] G. Apostolopoulos, D. Williams, S. Kamat, R. Guerin, A. Orda, and T. Przygienda, "QoS Routing Mechanisms and OSPF Extensions," RFC 2676, August 1999

[24] D. Katz, K. Kompella, and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2," RFC 3630, September 2003

[25] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1, Functional Specification," RFC 2205, September 1997.

[26] D. Awduche, L. Berger, D. Gan,T. Li, V. Srinivasan, and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels," RFC 3209, December 2001

[27] L. Andersson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas, "LDP Specification," RFC 3036, January 2001

[28] T. Kimura, S. Kamei, T Okamoto, "Evaluation of DiffServ-aware constraint-based routing schemes for multiprotocol label switching networks," Networks, 2002. ICON 2002. 10th IEEE International Conference on, 27-30 Aug. 2002 Page(s): 181 –185

[29] V. Fineberg, C. Chen, X. Xiao, "An end-to-end QoS architecture with the MPLS-based core," IP Operations and Management, 2002 IEEE Workshop on, 2002 Page(s): 26 –30

[30] E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture," RFC 3031

[31] Faloutsos, M., Faloutsos, P., and Faloutsos, C. "On Power-Law Relationships of the InternetTopology". In ACM SIGCOMM 99. September 1999.

[32] Blake, S., Black, D., Carlson, M.,Davies, E., Wang, Z., And Weiss, W. "An Architecture for Differentiated Services". RFC 2475. December 1998.

[33] Nichols, K., Blake, S., Baker, F., And Black, D. "Defintion of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers". RFC 2474. December 1998.

[34] Heinanen, J., Baker, F., Weiss, W., And Wroclawski, J. Assured "Forwarding PHB Group". RFC 2597. June 1999.

[35] Jacobson, V., Nichols, K., And Poduri, K. "An Expedited Forwarding PHB". RFC 2598. June 1999.

[36] CSELT DPC 1999.03108, "Telefonia su IP: problematiche e soluzioni per una rete di transito in tecnologia IP".

[37] P. Senesi, P. Ferrabone, G. Gritella, R. Rinaldi, M. Siviero, "Telephony over IP: theoretical modelling and lab experiments", [Universal Multiservice Networks, 2000.

[38] S. Floyd. "Comments on Measurement-based AdmissionsControl for Controlled-Load Service". Computer Communication Review, 1996.

[39] L.Alcuri, F. Saitta, "Telephony over IP: A QoS Measurement-Based End to End Control Algorithm", CCCT 2003 Orlando (FL), August 2003

[40] Handley M., Schulzrinne H., Schooler E., Rosenberg J., "SIP: Session Initiation Protocol", RFC 2543. IETF. March 1999.

[41] A. Vemuri, J. Peterson, "SIP for Telephones (SIP-T): Context and Architectures", draft-ietf-sipping-sipt-04.IETF. June 2002. Best Current Pactice

[42] Best Current Practice for Telephony Interworking", draft-zimmerer-sip-bcp-t-00.txt . IETF. April 2000. Informational

[43] J. Elwell, F. Derks, P. Mourot, O. Rousseau, "Interworking between SIP and QSIG", draft-elwell-sipping-qsig2sip-03.txt.IETF. October 2002. Internet-Draft

[44] E. Zimmerer, J. Peterson, A. Vemuri, L. Ong, F. Audet, "MIME media types for ISUP and QSIG Objects", Request for Comments: 3204. IETF. December 2001.Standards Track

[45] M. Arango, A. Dugan, I. Elliott, C. Huitema, S. Pickett, "Media Gateway Control Protocol (MGCP Request for Comments: 2705.IETF. October 1999