Title:	Document Version:
D2 **QoS Roadmap**	1.0

Project Number:	Project Acronym:	Project Title:
IST-2000-26418	NGNI	Next Generation Networsk Initiative

Contractual Delivery Date:	Actual Delivery Date:	Deliverable Type* - Security**:
30/09/2002	08/10/2002	R – PP

* Type: P - Prototype, R - Report, D - Demonstrator, O - Other
** Security Class: PU- Public, PP – Restricted to other programme participants (including the Commission), RE – Restricted to a group defined by the consortium (including the Commission), CO – Confidential, only for members of the consortium (including the Commission)

Responsible:	Organization:	Contributing WP:
Maurizio Molina	(NEC-Europe Ltd)	WP3

Authors (organizations):

Ilka Miloucheva (Salzburg Research mbH), Sandra Tartarelli (NEC-Europe Ltd)

Abstract:

In the recent past research and development efforts in the area of technologies for QoS support mainly concentrated on the definition of architectures and protocols for differentiated packet treatment at the IP level. However, their wide deployment is still lacking as a result of lack of general and simple rules for their configuration. But even if they were, would this really bring to ubiquitous, stable end-to-end QoS in NGN networks? It may not be the case, given that today "QoS is not there" even if most of the major operator's and ISPs backbone links are operated (on the average) at a very low utilization level. In this document we put in evidence that QoS has several aspects, not all (and not necessarily) related to the IP level packet treatment, and we outline the specific areas where there's need to concentrate research and engineering efforts in order to improve today's user perception of QoS.

Keywords:

Access Networks, Availability, Quality of Service, Security, Traffic Engineering

QoS Roadmap

EDITOR: Maurizio Molina (NEC)

AUTHORS: Maurizio Molina, Sandra Tartarelli (NEC)
Ilka Miloucheva (SR)

1. Introduction

In the recent past research and development efforts in the area of technologies for Quality of Service support in Next Generation Networks mainly concentrated on the definition of architectures and protocols for differentiated packet treatment at the IP level. Architectural frameworks like Intserv, Diffserv and MPLS have been standardized, as well as the accompanying concrete mechanisms and signalling protocol to support them (RSVP, CR-LDP, RSVP-TE, definition of EF and AF Per Hop Behaviour for Diffserv).
Standards in this area are mature enough, and implementations in commercial products are widely available. In spite of this, there's a general feeling in the networking community that these mechanisms are too complicate to be operated in practice and that the improvement that can really bring to end user's experienced QoS is limited.
This feeling, coupled with the observation that most of the major operator's and ISPs backbone links are still operated (on the average) at a very low utilization level, can lead to the belief that networks are good enough and that there's no need for further research for QoS support in NGN.

Thousands of researchers in the area, having spent years of effort in studying and simulating packet scheduling algorithms in Routers or ATM switches, are certainly ready to question this statement on the basis that it's "brutal" and not completely true, and that things may change in the future when the next killer applications will be found, etc.
But from an industrial and business perspective this doesn't make sense. One should more honestly admit that it's true that most of the network backbones are today "good enough". Some of the above-mentioned QoS support mechanisms have in fact been thought and developed under the assumption that the long distance bandwidth and the router processing speed would be the main Internet bottlenecks. But this is already history, as the industry replied by developing new hardware switching architectures or by adapting existing ones (e.g. ATM) to the IP level, while bandwidth availability on optical fibres has continued to increase more than the demand.

Nevertheless, we share the view that the "no QoS need" statement *is* actually "brutal" and not completely true, but it has to be questioned with different counter-statements. And this document mainly concentrates on giving them.

First, there's no need to wait for the next killer application. End to end QoS is today *not* there ("Not even Undergrads would use VoIP" [1]) or not always there, which is equally bad ("No one wants unpredictable, poor performance [1]"). Also, many Service Providers see QoS as a differentiation mean and revenue opportunity, and this will justify investments needed to provide QoS [2].

The most obvious reason why QoS is not ubiquitously there today is because there's a lack of its support in the access networks. For Wireless access, it's quite obvious that the growth of available bandwidth will still be limited for long time, so mechanisms to differentiate and protect real time, QoS demanding traffic from bursty data traffic will be more and more necessary. Always in Wireless networks, because of the race towards the offering of higher bit rate services that pushes for smaller and smaller cell sizes (micro and pico cellular environments), the issue of efficient and scalable mechanisms for handoff support becomes more relevant. In Wireless LAN protocols, the main issue is in the introduction of layer 2 QoS mechanisms.

In optical based Metro access networks the two today's dominating technologies are Gbit Ethernet and ATM over Sonet/SDH. Both were quite successful, but have main disadvantages in the lack of a reliable resilience methodology (Ethernet) and in the very inefficient utilization of resources (ATM over Sonet/SDH). Resilient Packet Ring (RPR) is a new standard from IEEE (802.17) that promises to join high bandwidth availability, resource utilization efficiency and a "native" methodology for network resilience.
Also x-DSL access introduces new challenges for QoS support. The issue of the QoS support in access networks is dealt in section 2.

Another limiting factor to the availability of good, ubiquitous, End-to-End QoS is terminal performance. As regards Voice/Video/multimedia terminals, still a lot of effort is needed to improve codec performance. In particular, coding schemes explicitly designed for a packet environment and able to adapt to a variable quality in the downstream network may gain more and more momentum.
On the server side (e.g. speaking about servers supporting heavily visited web sites), a big improvement in performance can be achieved by moving the servers in dedicated, professionally operated Internet Data Centres (IDCs), potentially replicated in multiple location all over the world. Nevertheless, there are still a lot of open points related to how to intelligently select the nearest IDC location, how improve the efficiency of these shared server farms having to deal with extremely variable and unpredictable loads, and how to prioritise HTTP requests once the final server has been chosen. Aspects relative to terminal and server performance are detailed in section 3.

Always in the direction of giving QoS a broader meaning than the simple "IP packet treatment", one must emphasize that besides perceiving good quality when a communication is established, QoS also means to always find the network and the end service up and running. In other words, service "availability" in a broad sense plays a fundamental role in user's perceived QoS. We identify at least two perspectives from which to watch at it.
The first one is service survivability to node and link faults. Nowadays, survivability is often implemented (independently) both at the SDH (SONET) level and at the IP level. Service recovery at the SDH level is fast and reliable, but requires a complete duplication of physical resources, while recovery at the IP level (with routing protocols finding alternative paths) doesn't require additional resources but has convergence times on the order of minutes (unacceptable for paying, real-time services). An alternative that is gaining momentum is to implement the protection at the MPLS layer, where it would be possible to avoid the duplication of resources dedicated to protection (backup path sharing) and at the same time meet the stringent requirements on quick restoration times needed for high quality services. Furthermore, by protecting at the MPLS layer it's possible to selectively protect some flows (paying more for that) and leave others unprotected (or protected with fewer guarantees). ISPs and NSPs look at the possibility to offer connectivity services with different degrees of protection guarantees and recovery times as a way to differentiate themselves and better target the different categories of customers.
The second aspect of availability, strictly related with security, is the protection against DDoS attacks, whose effect, from an end user perspective, is indistinguishable from poor QoS due to congestion. Defence from DDoS attacks requires actions like securing servers, deploying firewalls, and filter potentially "foe" packets on network routers. Very important is that these activities are carried out in parallel by all (or the majority) of the entities administering Internet's domains: the nature of DDoS attacks is in fact such that the local security of a host or of a router impacts on the security of everybody else.
The detailed discussion on forthcoming hot issues in availability is in section 4.

As previously mentioned, one of the main arguments brought by "QoS sceptics" is that networks today are good enough. As noted in [1], it's certainly true that the need of applying complicate schemes for QoS support in the core is today limited by the slow rate in the access, that smoothes the traffic burstiness. However, this may become less and less true in the future, if end to end Gbit LAN services will eventually allow users to "really" burst at Mbit/s speeds or higher ("Next Generation VPN"). In such a scenario, it may make sense to widely apply Traffic Engineering (TE) methodologies to avoid sudden network congestions and resource under-utilization. How to apply Traffic Engineering rules exploiting MPLS and the Diffserv paradigm is described in 5. Note that fundamental for applying TE techniques is the availability of timely and accurate measurements giving the operator a reliable picture of the status of the network (e.g. load) and possibly of the QoS delivered by its own network. Being QoS an end-to-end matter, it's also becoming more and more important to share these measurements between operators in a coordinated way. An overview of a recently established IST project (INTERMON) performing research exactly in this area is given in 5.1.1. Always in section 5 we also introduce the work recently undertaken by the IETF WG NSIS (Next Steps In Signalling) aimed at defining the requirement for a new (or modified) signalling protocol for requesting QoS in the Internet. Always in the same section we describe another interesting area of active research, which aims at introducing QoS in the Internet by evolving (not changing) the current Best Effort Paradigm.

In section 6 we give an overview of the areas on which the networking research communities are currently mostly focusing, by listing and grouping the topics of the Call for Papers of the most recent and major international conferences in teletraffic and communications.

Finally, we conclude this document with a quite detailed review of the activities in the area of QoS support performed by recently concluded or ongoing IST projects or other publicly founded projects in Europe, outlining for some of them, what future research activities may derive from their achievements. All this is contained in section 7.

2. QoS in Access networks

As mentioned in the introduction, it is foreseeable that in the near future QoS will be primarily a problem in the access networks. In wireless and x-DSL access networks, where that bandwidth limitation is still an issue, the main challenge is how to protect traffic with stringent real time requirements (like VoIP or video conferencing) from bandwidth greedy, non real rime traffic like data traffic or video streaming. In metro area optical access networks (that in most cases will be the network gathering the traffic generated on wireless or x-DSL) the main issue is on the contrary to have a reliable mechanism to protect all the gathered traffic from fibre or node failure.

2.1 Wireless networks

2.1.1 QoS differentiation in the air interface of UMTS Networks

In order to accommodate the envisaged different types of services, ranging from voice, to real time video, to video streaming and data, UMTS standards define 4 different QoS classes, that differ on their ability to guarantee bounded delays, different BER levels, bandwidth guarantees[12]. These 4 classes are called "Conversational", "Streaming", "Interactive" and Background. Possible associated applications are, respectively, VoIP, Audio/Video, Web Browsing and e-mail. UMTS standards introduce also the concept of "bearers" that are logical transport entities that must support traffic delivery with one of the above-mentioned classes. QoS has of course an end-to-end scope, but the supporting bearer is decomposed in a sequence of bearers (one per each portion of an end to end communication) each having to meet some requirement in order to support the desired end-to-end QoS. The main challenge is to provide a quality support (and differentiation) in the Radio bearer service, spanning over the Radio Access Network (RAN).
In the UMTS the RAN is composed by two types of nodes: the Base Stations – BS - (interfacing with the Mobile Stations- MS) and the Radio Network Controllers – RNC - interfacing with the BS on one side and with the UMTS core network on the other. The RNC is responsible for all mobility related aspects, while the BS plays a major role for what concerns resource management on the air interface.
[13] proposes a class based scheduling algorithm to support class differentiation over a CDMA air interface (which will be the dominant air interface in 3G systems). The classes basically differ for how "robustly" they preserve their negotiated bandwidth in case of congestion that is likely to happen on the air interface due to the following reasons: burstiness of user traffic, deterioration of channel condition that leads to the need of increasing the Signal to Interference ratio (SIR), decrease of capacity due to increased interference from other cells, unconstrained user mobility. When such conditions are detected, the rates of communications ongoing in the cell (controlled by the BS) are selectively reduced on the basis of the "elasticity" of the respective class. One can envisage, for example that the applications of a "Platinum" class won't have their rate reduced at all, while those of a "Gold" class will face some reduction, those of a "Silver" class an even stronger one. The transmission rate is reduced by changing the spreading factor (a property of CDMA based systems). In the proposed framework, the intelligence to actuate such a mechanism resides into the BS, which determines the transmission rate for each active CDMA code on the next frame. This requires provision for fast control messaging over the air interface.
As clarified in [14], the approach proposed in [13] belongs to the category of "pure rate based" methods that control the access to the limited cell bandwidth. There are also "pure

power based" methods that simply act on the transmit power of the mobile station and of the BS. An interesting novel approach, proposed in [14], is to control in a combined way power and rate. The authors show that such a method, based on a genetic algorithm (GAME: Genetic Algorithm for Mobile Equilibrium) can outperform both purely power based and control based methods (the challenge laying in containing the complexity of the genetic algorithm computation).

2.1.2 Wireless LAN

The IEEE 802.11 Wireless LAN standard[36] has been widely accepted in many different environments today [37]. In its current version, the 802.11 standard can be considered as an extension to Ethernet, which supports only Best-Effort services. Recently, the interest in wireless networks supporting Quality-of-Service (QoS) is growing tremendously [38], [39], [40], [41], [42], [43], [44], and the IEEE 802.11 Working Group has established an activity to enhance the current 802.11 MAC (Medium Access Control) protocol to support applications with QoS requirements.
The discussion of the schemes for QoS support in 802.11 is in its way of being completed by the IEEE 802.11 Task Group E (TGe). The latest draft[45] by TGe. 802.11e specifies the 802.11e supplement standard and introduces two access mechanisms: the Enhanced Distributed Coordination Function (EDCF) and the Hybrid Coordination Function (HCF). The challenge with 802.11e lies in determining how to configure the EDCF and the HCF to provide the desired services. Providing guaranteed throughputs is widely accepted as one of the desirable services in a QoS architecture [46]. This service suits well, e.g., the needs of data communications. In line with the above considerations, an open issue is how to provide throughput guarantee services in an 802.11e Wireless LAN under the EDCF.

2.1.3 Support of frequent handoffs in micro and pico cellular environments

To support higher data rates, 3G wireless access networks will have more and more to rely on micro and pico cellular environments. This increases the frequency of handoffs due to cell change and puts new challenges on the design of algorithms that can efficiently support mobility while still keeping the mobility related information exchange between terminals and the BS at a reasonable rate. Mobile IP, specified in RFC 2002, supports mobility with the concept of a Home Agent exchanging location information with a Foreign Agent whenever the user moves between networks. Simply extending it to support the mobility between cells wouldn't scale. To overcome this, several proposals have been submitted to the IETF [15], [16]. Common to them is the concept of grouping cells into "registration domains" and avoiding to send to Home and Foreign Agents location updates when the mobility is within a single registration domain. In addition to the issue of limiting the mobility related signalling, the increased handoff frequency brings the problem of reserving in advance some resources for handoff calls. In a scenario where the handoff within a call is the rule (rather than the exception, as in 2G networks) it is necessary to design mechanisms to treat handoff calls differently from new calls (it is much more annoying to experience a call drop once the communication is established rather than getting a busy signal before the establishment). [17] Gives an architectural overview of the problem and proposes to reserve a percentage of the resources per each cell (guard channels) for handoff calls. Some results are presented, assuming uniform load distribution of calls and users within a registration domain and simple models for user mobility and call arrival/termination process. The problem with the dimensioning of the guard channel is of course the trade off between a reduced handoff call drop probability and an incremented new call blocking probability. Another interesting approach for dynamically dimensioning the guard channel, without having to rely on any

trivial (and potential inaccurate) modelling of load distribution and user mobility behaviour is reported in [18]. The approach is based on predicting future needs for handoff calls from recent load measurements.

2.2 xDSL

xDSL technology has greatly increased the Internet access speed for residential and small corporate subscribers. However, xDSL gives also the potential of carrying in parallel, on the same physical medium, data and voice traffic. An xDSL user can still browse the internet or download mails while making a phone call, either from a legacy phone connected to the POTS port of the ADSL modem or from a softphone on the PC itself, of from a new generation IP phone. Also, with PPP over Ethernet (rfc 2516) a whole group of PCs can share the same DSL modem.

This potential convergence of Voice and data communications in a single access point (the DSL modem) introduces however the challenge to guarantee that in every condition the real time traffic is protected from non real time, bursty data traffic.

The challenges are both at the terminal/modem level and at the network level. To understand why, let's recall that the most common today's xDSL deployment are based on ADSL technology (thus, depending on the commercial offering, with speeds around 128Kbps upstream and 640Kbps downstream) and on ATM encapsulation, i.e. there's an ATM Permanent Virtual Circuit (PVC) from the ADSL modem that passes through a DSLAM and terminates typically at the edge of an ISP or corporate IP network, in a so-called "Broadband Access Server". From the DSLAM to the IP network the user's ATM PVC is typically carried together with a lot of others in an ATM Permanent Virtual Path (PVP). For a correct and stringent QoS guarantee to data traffic, the optimal choice would be to have two ATM PVC from the ADSL modem, one fed by data traffic and the other by real time voice traffic. This is currently not supported by the majority of ADSL modems, either Ethernet or USB based. Note that in the absence of a VC dedicated to carry real-time traffic, even a single upstream data packet of 1500 bytes, at 128kbps, could block a voice packet for 90ms, which is unacceptable.

Even if the problem of providing a separate VC for voice (or, in the near future, also video) traffic with real time requirement can be solved, there's then the problem of providing the bandwidth for this PVC from the DSLAM up to the IP backbone network (i.e. in the ATM access network, where realistically bandwidth is still a precious resource). The ATM PVC is normally bundled together with a lot of others in an ATM PVP, and normally only the bandwidth of the PVP is guaranteed. Then, not knowing if the user is really transmitting traffic on its PVC, and with which type of audio/video codec, the only way out to provide guarantees would be a high over provisioning.

One solution, envisaged by several operators [23] [24], is to introduce an intermediate functionality between the "application" layer processing the user request for call establishment and the network layer simply relaying the data traffic. This functionality can be viewed as a Network Resource Management system, or as a "client" of a more general Network Resource Management system. The main role of it is to control the availability of resources in the PVP (based on the knowledge of the PVP bandwidth itself and the number and bandwidth of ongoing voice or video calls carried on it), blocking calls (i.e. the signalling SIP or H.323 packets) if not enough resources are there. Additionally, this functionality may have the intelligence to request to the Network Resource Management system the resizing of the PVP, for resource optimisation. Also, the same functionality should "instruct" the edge device of the IP network to open a "pinhole" for letting the related voice or video traffic flow through the IP backbone network. This pinhole opening functionality, also referred as

Middlebox control, is currently studied in the IETF WG MIDCOM and it's very important for security reasons and for preventing theft of service.
Note that the above-described issue is still there even if the access ATM network is replaced by a pure IP tunnelling or by MPLS.

2.3 Optical access Networks

Sonet/SDH is currently the dominating technology in the Metro access area. It is typically deployed in physical ring topology, and has fast and reliable mechanism to protect the ring from fiber cuts or node failures. Historically, ATM has been inserted as an intermediate layer between the circuit-oriented Sonet/SDH and IP. This solution, however, is regarded by many operators as being extremely inefficient, due to the presence of a whole additional layer (ATM), with all the related protocol inefficiencies and management problems. The simple elimination of ATM (IP over Sonet/SDH) doesn't solve all the issues, as then any logical network topology creation (which is the main feature provided by ATM) has to be realized with an extremely coarse granularity, dictated by the Sonet/SDH interface speed hierarchy and with great amounts of wasted bandwidth.

Ethernet has evolved in the past years from a simple Local Area, low bandwidth, shared access network technology to a point to point technology based on optical fibres that can provide bandwidths up to 1 Gbit/s (and soon 10Gbit/s) on spans over 50 miles. It's therefore in principle suitable also as a Metro Access network technology. Also, its packet nature makes it possible to avoid multiple intermediate layers with IP packets being directly encapsulated in Ethernet frames. However, simply extending Gbit Ethernet to work on the currently widely deployed optical rings in the Metro area isn't possible, as a ring is a shared media and an Ethernet based MAC access protocol cannot efficiently work at Gbit/s speeds and with Metro wide spans. Even so, Gbit Ethernet wouldn't have any built in protection mechanism, fundamental for network reliability.

Resilient Packet Ring (RPR) [11], being standardized by the IEEE 802.17 RPR Working Group, is an emerging solution that can have an efficient support for Ring topology and fast recovery from failures and at the same time retain all the inherent advantages of a packet based transport mechanism like Ethernet. Also it addresses features like fairness and congestion control, not present in the mentioned legacy technologies.
An RPR station is positioned on a double uni-directional link and performs three basic operations: adds packet (on one of the two directions) extracts packets having as a destination the station itself and otherwise forwards any other packet. The simplicity of these operations (no routing is needed, as packets will always reach the destination a ring topology) allows the interfaces to easily scale at high speeds. Also, the RPR stations monitor the ring occupancy and forward this information to all the other stations on the ring. Stations noting or being informed of congestion in some portions of the ring will limit their insertion rate according to a fairness algorithm. In case of a fiber cut everything continues to work in the same way provided that the two stations at the end of the failed link redirect all the packets destined to the filed link on the other direction of the ring. Target protection times are below 50ms.
Traffic feeders of RPR stations can be ordinary Gbit Ethernet interfaces. Note also that 802.17 concentrates only on the MAC layer, allowing its operation over existing physical layers (e.g. Sonet/SDH, Ethernet).
The main research issue related to RPR regards how to achieve ring-wide QoS/fairness goals

3. Terminal and Server performance

3.1 Improving VoIP terminal performance

Being QoS an end-to-end matter, the performances of the terminal equipment is integral part of the user's perceived QoS. In particular, three elements of VoIP terminal equipments can have a dramatic influence: the codec, the packetization scheme, the de-jitter (playout) buffer.
According to [19] and[21], the codecs mainly used in VoIP today (G.729, G.723.1, GSM-EFR, 3GPP AMR) were designed for traditional circuit switched networks, and their common weak point is that they are all based on the Code Excited Linear Prediction (CELP) paradigm. This paradigm makes all these codecs to be stateful, and thus "sensible" to packet losses or excessive delays, both very common in packet networks. [19] proposes, for example a software solution (NetEQ^{TM}) that can be based on a newly developed, proprietary codec called iLBC (recently also submitted to the IETF as a draft proposal [20]). iLBC is claimed to have much better performance than existing codecs in presence of both moderate and heavy packet losses.
The packetization scheme can also severely impact the quality of a VoIP call, because of the added delay it introduces. [22] recommends to avoid inserting multiple frames per IP packet.
The de-jitter buffer size is currently statically set in the majority of terminals, and this leads to a permanent, unnecessary additional delay. [19]. The software solution (NetEQ^{TM}) offered by [19] is on the contrary based on a dynamically adjustable de-jitter buffer. Also [22] supports the need of the dynamic adjustment of de-jitter buffers, and wishes that the parameters of de-jitter buffers can be exported (in a standard way) to network operators, so that they could optimise the network transmission delay on the basis of them.

3.2 Improving WEB Server Performance

The era of commercial exploitation of the Internet has already begun since several years. As more and more companies are offering their services over the web, it becomes fundamental not only *what* is offered, but also *how* it is offered. That is, whatever e-commerce site has nowadays to be professionally designed and with a captivating (and working) graphical interface. Fundamental is also that the site is always available and that it responds quickly to user's request. To address these reliability and responsiveness requirements, large Internet Data Centres were hundredths or thousands of servers host a lot of web sites have been created. IDCs basically solved the availability problems (they are hosted in protected buildings, with 24/7 surveillance, Power Continuity groups, batteries of firewalls usually protect the servers from DoS attacks) and the responsiveness component dependent from downstream bandwidth availability (IDCs are attached to the Internet backbone with huge pipes, rarely overloaded).
Still, a major component of server responsiveness to user's requests depends on server performance. The control of server response time is not trivial as usually very simple requests may trigger a complicate procedure before the response is generated (e.g. back-end databases may be fetched for user authentication, web pages may by customized and dynamically generated, etc.). Research activities in this area are still a hot issue, and they may be categorized as follows: server load balancing and migration, server assignment and caching, application level scheduling in web servers.

As said before, IDCs can host hundredths or thousands of hosts. Typically, the software implementing a web site resides on several of them, both for reliability and load balancing.

Dedicated appliances at the entrance point of the IDC typically perform the load balancing among all the web servers equally capable of responding to a single HTTP request. These appliances differentiate themselves for the Hardware or Software architecture, for the layer at which they perform the load balancing (e.g. per TCP connection, or per HTTP request, etc.) and for the specific load balancing algorithms they can implement (e.g. open loop vs. reactive). See [3] for a classification of the load balancing methodologies.

Although load balancing the HTTP requests towards a pool of servers hosting a specific site is a significant step towards performance improvement, the dimensioning of the number of servers per web site is still a challenging. With whatever static configuration, it's still quite easy to face anyway congestion problems, due to the highly variable and unpredictable load of requests towards web sites, that may be influenced by unpredictable events like success of advertisement campaigns, social, politic and natural events, sudden drifts of fashion influencing customers' behaviour. [4] propose a methodology for dynamically migrate servers from shared server pools or away from lightly loaded web sites to temporarily overloaded web sites.

For popular and world wide accessed web sites, the corresponding web server (or web server pool hosted in an IDC) isn't typically unique. More likely, the same server or server pool is spread over several physical locations all over the world. Established companies like Akamai [5] Exodus [6] or Mirror Image [7] offer a distributed server placement in different IDCs. Then they typically implement proprietary techniques to direct a client request to the best geographical location. [10] describes the challenges related to the task of clustering client requests towards the "closest server" and describe a solution (Webmapper) that can be coupled to an authoritative DNS server to perform this task. The method to build the metric for choosing the closest server is bases on server's passive monitoring of its incoming TCP connections.

[8] and [9] both address the problem of admission control and scheduling of HTTP requests in a WEB server. They introduce the concept of a "Session", which is closely related to some high level human activity. E.g. a session can be the whole set of activities related to the purchase of some goods, from the catalogue browsing to the filling of the shopping cart to the checkout. A session is of course composed of several, subsequent HTTP requests that can be correlated by means of the exchanged cookies (that are session-related identifiers carried in HTTP messages) or by the presence of an HTTP/1.1 persistent connection. An overloaded web server may decide to reject HTTP requests opening new sessions giving priority to the completion of ongoing ones.

4. Service Availability

4.1 Survivability to faults

Nowadays the survivability to faults (node or fiber failure) is normally implemented independently at the IP layer (rerouting) and at physical Sonet/SDH layer. IP layer protection has the advantage that network resources for protection need not to be dedicated, but the converegence times may be of the order of minutes, thus too slow for critical services. Sonet/SDH layer protection has dramatically lower restoration times (few milliseconds) but requires the complete duplication of physical resources. Both methodologies are completely "blind", with respect to the actual protection requirement of the IP flow affected from a network fault.
On the contrary, the connection oriented service provided by MPLS can be associated with several path protection/recovery mechanisms so that each flow is protected to a level adequate to its needs (and pays accordingly for that). The protection at this level of course eliminates the need of a protection at the physical layer.
[28] describes the issue and introduces a "Resilience-Differentiated QoS – RD QoS" architecture. First, four different RD-QoS classes are introduced. The classes differ for their resilience requirement (i.e. from "high" to "none") and for their needed recovery times (i.e. from "<100ms" to "not applicable"). Then, it's explained what extensions would be needed for existing QoS architectures (Intserv, Diffserv, MPLS) and associated signalling protocols to support the RD-QoS concept.
For Intserv, it's proposed to combine the three existing QoS classes with a two-bit attribute in the RSpec message signalling the Resilience class of the flow. Backup resources for flows demanding high resilience are reserved in parallel to primary resources. Of course, the Traffic Engineering extensions of RSVP need to be used to avoid reserving backup resources on the same set of links of primary paths. For Diffserv, "the bit patterns for resilience DSCP may either be taken from the DSCP standardized pool or the pool for local and experimental use". The Network Management reserves adequate resources "according to the estimated or negotiated (by SLA) amount of traffic having resilience requirements".
As regards MPLS, the "classical" MPLS Traffic Engineering functionalities (that can be, for example, implemented in a traffic engineering server) need to be extended to take into account the resilience requirement signalled by MPLS (MPLS signalling protocols, i.e. RSVP-TE and CR-LDP are already extended to support resilience signalling and link/node failure notification).

4.2 Protection from DDoS Attack

Successful Distributed Denial of Service (DDoS) attacks against popular web sites (Yahoo, Buy.com, Amazon, Datek, CNN and others) in the early 2000 and the echo this event had on the mass media made even the non-experts aware of the problem of Internet security. Among all the security flaws of the Internet, DoS attacks towards WEB or DNS servers are those that, from a user's perspective, result more similar to poor performance (or unavailability) of the server themselves and are therefore covered here.

A DDoS attack is based on the principle of an attacker machine that manages to install on a multitude of agent machines attacking software. Once the set of agent machine is sufficiently large, they start sending packets towards a victim server (or towards a set of victim servers). The servers suffer from the attack either because these packets are simply flooding their

network connections (brute force attacks, like UDP or ICMP flood) or because they exploit standard protocol characteristics that make the servers quickly run out of CPU or memory resources (e.g. SYN flood attack, CGI request attack).
One of the characteristics of the DDoS attacks is that it is difficult (or useless) to trace back the source of the attack. Difficult, because attacking software usually doesn't send packets with the real IP address of the agent machine, but rather with randomly forged ones. Useless, because the tracking of an agent machine, which may involve a lot of effort, would result in eliminating only one out of hundredths (or thousands) of agent machines. A careful attacker also will care of not directly controlling the attacked machine with ordinary network connections (otherwise, he would risk to be discovered) but e.g. through Internet Relay Chat channels, which are part of a legal and popular service and guarantee anonymity.

Being the attacker trace back difficult or useless, the defence from DDoS attacks usually concentrates on mitigating the impact of the attack itself. This can be achieved in three different ways: securing the server machines, placing firewalls in front of the servers and activating DDoS protection mechanisms on them, enabling defence strategies on the Internet Routers (particularly at the edge).

Securing the server machines involves a broad range of activities including applying known Operating System patches, closing some network ports, changing the protocol stacks to mitigate the effect of protocol attacks. A (beneficial) side effect of this is that the server itself becomes less easily an agent for attacking other machines (Internet Security, to this extent, is unfortunately highly dependent from how secure the *whole* Internet is….).

Today's firewalls (either hardware appliances or purely software solutions) aren't any more simply packet filters. Most of them have sophisticated solutions also to limit the negative effects of DDoS attacks on servers. For example, they can send fake SYN ACKs to clients in order to protect the server from SYN belonging to a SYN flood attack. While having a good firewall solution in front of a server farm is fundamental for enhancing security, it doesn't solve all security related problems. For instance, it doesn't protect from a brute force attack managing to flood the incoming network link. Also, due to the increasing level of complexity of firewall operations (e.g. the need of keeping per connection states), one must ensure that the firewall processing capacity itself doesn't become a system bottleneck. Recently, the concept of multiple firewalls with load balancing appliances in front of them has been introduced. Commercial products implementing this feature start to be available [25].

Implementing on Internet Routers mechanisms to fight DDoS attacks is finally the only way to avoid that attacking traffic reaches (or massively reaches) the victim servers. For example, we mentioned before that attacking software makes the agent machines produce packets with randomly forged IP addresses. Edge Routers can be aware of the client subnetworks connected to them, and should therefore discard packets coming from unknown subnetworks, that most probably are forged attacking packets. Another example is disallowing packets directed to broadcast addresses of the domain the Edge Router belongs to. Normally these packets are ICMP echo request packets with as source IP address the forged IP address of a victim server. If these packets go through and reach the broadcast addresses of the domain they get replicated and reach all the machines of the domain, which in turn will send an ICMP echo reply to the victim flooding it.

The refinement of DDoS attack and the response methodologies are continuing to evolve, sometimes so rapidly that's difficult to predict what he future areas of activity will be. What's particularly important, however, is that people involved in network security perceive more and more that cooperation is fundamental for achieving success [27]. As an example, a few

years ago, in response to e-mail spammers, system administrators started to modify their e-mail server software to prevent open relaying (i.e. transmission of e-mail messages not originating from or destined to their network). The next step was to compile and publish black lists of the remaining open e-mail relays. Many Internet providers subscribed to these blacklists and rejected e-mails coming from them. It's expected that for other security practices (e.g. ingress IP filtering) a similar course will be followed [26].

5. Evolving and Engineering today's QoS paradigms

5.1 Traffic Management / Traffic Engineering

In today's Internet, the management of IP networks relies almost exclusively on manual configuration of routers and on a limited amount of measurements data, resulting in a likely inefficient use of network resources. Moreover, current Interior Gateway Protocols (IGPs) use the shortest path to forward traffic. However, shortest paths from different sources overlap at some links, potentially causing congestion on those links. Besides, it might happen that the shortest path is highly congested, while a longer path between the same two routers is under-utilised. Traffic Engineering (TE) aims at optimising the traffic transport through an operational network, by attempting to minimize resource over-utilization whenever alternative resource is available in the network. A major objective of the Internet TE is the enhancement of network performance, both at the traffic and at the resource level.

In [29], the authors propose a process model to describe the most common activities carried out by a TE system. Such model consists of four phases:

- definition of the control policies that regulate the network operation. These policies might depend on different factors, e.g. the business model, the cost structure, the optimisation criteria etc;
- performance monitoring;
- network performance analysis and traffic characterization. This phase might aim both at preventing problems (e. g. points of failure) or at detecting the causes of existing problems in an operational network;
- performance optimisation, i.e. the choice of the most appropriate solutions among a set of alternatives.

5.1.1 Tools for TE

Although the need for a systematic TE approach in the management of large IP networks is commonly recognised, the networking industry is still lacking software systems that can support traffic measurements and modelling, essential for an effective TE.

A considerable contribution in this direction is represented by NetScope **[30**], a unified set of software tools for managing the performance of IP backbone networks. Usage and configuration data is extracted from the network elements and interpreted to build a global view of the network. The network provider can then infer and visualise the network-wide implications of local changes in traffic, configuration and control, by testing them in a simulated environment. A typical application would be the location of heavily loaded links, the identification of the responsible traffic demands and the re-configuration of intra-domain routing to reduce the detected congestion. The same authors explain in more detail in **[31**] some techniques to derive traffic demands from an operational network, while in **[32**], they discuss some guidelines to identify router configuration mistakes, with particular focus on intra-domain routing. The activity reported in the three mentioned papers is limited to intra-domain TE in a large IP backbone network (in the specific case the AT&T backbone). However, this effort clearly points out the increasing need of ISPs and network providers to rely on tools able to automatically react to rapid changes, congestion situations and other events that might affect the performance of their network.

Even more challenging is the task of designing a system capable of engineering traffic in an inter-domain scenario. An attempt in this direction is the focus of the IST INTERMON project[35]. The project aims at building a system that enables a variety of advanced inter-domain tasks, ranging from TE, network planning, SLA monitoring, accounting/billing. Further details on this project are given in 7.2.1.
The increasing interest of the Internet community towards measurement activities, especially as a support to network management and traffic engineering, is also witnessed by the work carried out by standardisation bodies, and in particular of the IETF. IPFIX ([53]) and PSAMP ([54]), are two recently established WGs dealing with the standardization of methods for measuring and reporting traffic at the flow level and sampling packets, respectively.
[55] is the Internet Measurement Research Group recently established by the Internet Research Task Force to perform basic research in the area.

5.1.2 MPLS Traffic Engineering (MPLS-TE)

Recent developments in multiprotocol label switching (MPLS) and differentiated services (DiffServ) have opened up new possibilities that extend the traditional functions of TE.
In particular, MPLS-TE facilitates the implementation of a number of functionalities, as compared to traditional TE. The key components of MPLS-TE are:

- path management, i.e. all aspects related to the selection and maintenance of routes, or more specifically LSPs (Label Switched Paths);
- traffic assignment, i.e. the allocation of traffic to a proper LSP;
- the distribution of network state and topology information;
- the network management, that determines the ease with which the network can be observed and controlled.

In **[33**], the authors present the design of the GlobalCenter' MPLS system (at the time the paper was written, GlobalCenter was one of the 10 largest ISPs in the US) and discuss practical issues of TE and techniques to provide QoS in an MLPS network. The design and implementation of a software system (RATES) that supports traffic engineering in MPLS networks is also described in **[34**].

Mitra and Ramakrishnan present in **[48**] techniques for traffic engineering in QoS-supported data networks, in particular MPLS or MPLS plus DiffServ networks. These techniques use primitives, typically employed for Linear Programming problems, which allow achieving the required speed of response and scalability. To handle the end-to-end constraints on routing imposed by QoS considerations, the authors introduce the concept of *admissible route sets*, $R(s,s)$, which are specific to each QoS service class, s, and (source, destination) pair s. For instance, real-time services, such as voice and video, may require routes with limited length (to avoid large propagation delays) and small number of hops. The techniques presented handle routing constraints embedded in the admissible route sets. Techniques to handle priorities are also discussed.

In [49], a new preemption policy for LSPs is proposed and complemented with an adaptive scheme that aims to minimize rerouting. The preemption policy combines the three main optimisation criteria: number of LSPs to be preempted, priority of LSPs to be preempted, and amount of bandwidth to be preempted. The preemption policy is complemented by an adaptive scheme that selects LSPs with lower priority and reduces their rate in order to accommodate the new high-priority LSP setup request.

5.1.3 DiffServ TE

Traditionally, MPLS mechanisms operate on an aggregate level. However, in a DiffServ environment it is possible to enhance network performance and efficiency by performing TE at a per-class level (for instance by taking different actions respectively for EF, AF and BE traffic classes) instead of at an aggregate level. In order to operate TE at per-class level, every traffic trunk[1] in a given class is mapped on a separate LSP and this allows the trunk to follow paths that meet constraints specific for the given class. This is referred to as "DiffServ-aware Traffic Engineering (DS-TE)" **[47**].

Networks that would benefit from DS-TE are for instance:

- networks where bandwidth is scarce (e.g. access)
- networks where high priority traffic is significant compared to the link speed
- networks where the proportion of traffic classes is not uniform over the whole topology.

As an example we can consider large voice trunks. For delay/jitter reasons it is undesirable to carry more than a certain "moderate" percentage of EF traffic on any link. The rest of the available link bandwidth can be used to route other classes corresponding to delay/jitter insensitive traffic (e.g. Best Effort Internet traffic). During normal operations, the VoIP (Voice over IP) traffic should be able to pre-empt other classes of traffic (if these other classes are designated as pre-emptable and they have lower pre-emption priority), so that it will be able to use the shortest available path, only constrained by the maximum defined VoIP link utilization ratio/percentage.

In the above scenario, DS-TE allows also to enhance the rerouting function, in case link or node failure occurs. Indeed, with existing TE mechanisms it is possible to reroute high priority traffic separately from the other classes of traffic. However, in the considered example it is also required that the high priority traffic does not exceed a "moderate" percentage on each link. Therefore, high priority traffic should be rerouted to a new link only up to a certain percentage. After that, a second link should be selected and filled again only up to the "moderate" percentage and so on. The remaining capacity on each link should then be used by other traffic classes. This is not achievable using standard TE mechanisms.

5.1.4 SLA Management

A Service Level Agreement (SLA) is a contract between a network provider and a customer that defines the service to be provided. Typical aspects covered by SLAs are: availability, performance and customer service. In IP networks that do not support Differentiated Services, IP SLAs are difficult to be defined and provided.

In [50], the authors discuss the current services offered by ISPs and investigate the limits imposed by traditional IP networks. The authors report that, currently, ISPs typically offer assurances only within their backbone, but not end-to-end. An exception to this is represented by WorldCom, that offers end-to-end SLAs for VPN services. However, the performance metrics are in this case based on very large time scales (one month is typical), the average

[1] In [29] a traffic trunk is defined as an aggregation of traffic flows belonging to the same class which are forwarded through a common path. A traffic trunk may be characterized by an ingress and an egress node, and a set of attributes which determine its behaviour in the network.

utilisation on the access link has to be less than 50% and only latency, but not packet loss is considered. It is clear that currently the end user is not offered a significant level of assurance.
[51] proposes a structure for QoS-centered SLAs, and a framework for their real time management in multiservice packet networks. The SLA is structured to be fair to both the service provider and their customer. An SLA monitoring scheme is presented in which revenue is generated by the admission of flows into the network, and penalty incurred when flows are lost in periods when the service provider is not SLA compliant. In the SLA management scheme proposed, the results of a prior off-line design are used, in conjunction with measurements taken locally at ingress nodes, to classify the loading status of routes. The effectiveness of SLA management is measured by the robustness in performance in the presence of substantial diversity in actual traffic conditions. The SLA considered in [51] are for QoS assured delivery of aggregate bandwidth from ingress to egress nodes; however, the control and signalling (i.e. route selection and admission control) is performed at at the granularity of flows or calls. The authors propose a measurement-based procedure for route selection that is implemented at ingress nodes and favours undersubscribed routes and previously successful routes. Moreover, admission control is performed in such a way that for undersubscribed routes acceptance is unconditionally accepted, while for oversubscribed routes the call can be accepted only provided that some conditions (explained in detail in the paper) are satisfied.
The differences that can exist between individual and aggregate loss guarantees in an environment where guarantees are only provided at an aggregate level are analysed in detail in [52]. The focus is on understanding which traffic parameters are responsible for inducing possible deviations and to what extent. In addition, the authors seek to evaluate the level of additional resources, e.g., bandwidth or buffer, required to ensure that *all* individual loss measures remain below their desired target. The paper's contributions are in developing analytical models that enable the evaluation of individual loss probabilities in settings where only aggregate losses are controlled, and in identifying traffic parameters that play a dominant role in causing differences between individual and aggregate losses. The latter allows the construction of guidelines identifying what kind of traffic can be multiplexed into a common service class.

5.2 Next Steps In Signalling – NSIS

Next Step in Signalling (nsis) [56] is an IETF working group focusing on signalling Quality of Service (QoS) over the Internet. The motivation for setting up the WG derives from the consideration that even if standard mechanism for supporting QoS in the Internet are consolidated, (Intserv, Diffserv) and implementable in single controlled domains, it is still very difficult to obtain end-to-end QoS (which is the only relevant one) in a composite multi administrative domain scenario. Among the ideas driving the work in progress we can mention the wish to decouple signalling from resource reservation (the latter being considered strictly a network management activity), the requirement that the users shouldn't be forced to understand how different domain allocate resources, the observation that some service providers will still want to use contractual means for reserving bandwidth rather than protocol means, the requirement that Application layer must wait until QoS requirements are fulfilled before setting up sessions.
Also, the WG looks for a modular protocol solution able to provide signalling in different part of the nework (end-to-end, edge-to edge, end-to edge), with different technologies, able to cope with mobility and mobile hosts and having a complete security support.
At this point in time, only requirements and frameworks for signalling have been discussed, future works will be to propose solutions meeting the requirements.

Particular attention has been put on carefully considering which portion of existing protocols could be reused. For instance, RSVP (which is the only standarized protocol to signal Qos requirements throughout Internet) is being considered, but separately from the Intserv framework for which it was originally developed. A modified version of RSVP could be the basis for the new signalling protocol.

5.3 Improving the best effort model

The Internet community has already been discussing for some time, whether it is worth extending the Internet architecture to provide quality of service guarantees or it is much simpler and more efficient to overprovision the network and keep the current best effort model. However, lately an alternative and intermediate position is becoming more and more popular, i.e. to improve the current best effort model.

In [57] Gevros et al. support the idea of using control structures to protect best effort traffic and even introduce some sort of service differentiation, but without "sacrificing the best effort nature of the Internet or stressing its architecture beyond its limits and original design principles". In [57], the authors revisit the best effort service model and the problem of congestion while focusing on the importance of cooperative resource sharing to the Internet's success, i.e. for instance, that all the users react by decreasing the rate if congestion occurs. Furthermore, the authors present a review of the congestion control principles and mechanisms, which facilitate Internet resource sharing.

TCP was originally designed exactly with the purpose of making all users sharing a link slowing down their transmission rate in case of congestion on that link. Since TCP was designed, however, the situation has significantly changed. With the increasing growth of non-TCP traffic (e.g. media streaming), congestion control had to be extended to non-TCP flows. A non-TCP flow not sending more than TCP flow under similar network conditions is said to be "TCP-friendly". However, making a non-TCP flow behave similarly to a TCP one is not straightforward. An overview of method for this purpose can be found in [58].

Another work that aims at improving the traditional best effort model is proposed in [59]. The purpose in this case is to seek to place more of the control in the hands of the end system or user, with simple functionality in the router. Using insights from economics and control theory, the authors show how cooperation between end systems and the network can be encouraged using a simple packet marking scheme. The network distributes congestion feedback information to users via packet marking at resources, and users react accordingly to obtain differential QoS. This approach contrasts with current proposals for creating differential QoS in the Internet, that typically rely on classifying packets into a number of classes with routers treating different classes accordingly. In the latter case it is the router, which plays a critical role in guaranteeing performance, while in the approach proposed in [59], the control is in the hands of the end system or user.

[60] also aims at retaining most of the aspects of the traditional best effort model, but at the same time seeks to achieve some service differentiation based on the type of application. To this end, the authors propose an alternative best effort (ABE) service, which relies on the idea of providing low delay at the expense of less throughput. The objective is to retain the simplicity of the original Internet single-class best-effort service while providing low delay to interactive adaptive applications. With ABE, every best effort packet is marked as either green or blue. Green packets are guaranteed a low bounded delay in every router. In exchange, green packets are more likely to be dropped (or marked using congestion

notification) during periods of congestion than blue packets. For every packet, the choice of color is made by the application based on the nature of its traffic and on global traffic conditions. Typically, an interactive application with real-time deadlines, such as audio, will mark most of its packets as green, as long as the network conditions offer large enough throughput. In contrast, an application that transfers binary data such as bulk data transfer will seek to minimize overall transfer time and send blue traffic. ABE is different from differentiated or integrated services in that neither packet color can be said to receive better treatment; thus, flat rate pricing may be maintained, and there is no need for reservations or defining traffic profiles.

6. Basic research

This section is meant to give an overview of the areas on which the networking research communities are currently mostly focusing. The following list of topics is extracted from the recent Call for Papers of the major international conferences in teletraffic and communications. One of the clear messages is that the success of a network is heavily dependent on its ability to offer *reliable* and economically advantageous service differentiation. Essential requirements are traffic management mechanisms capable of providing necessary QoS guarantees and traffic engineering procedures, which ensure that these mechanisms are used in a cost effective way.

6.1 Network architectures, technologies and services

- Next generation networks
- Broadband transport networks
- Access networks
- Wireless networks and wireless LANs
- Optical networks
- Wireless, mobile, and pervasive networking, networking with specialized devices and sensors
- Heterogeneous networks and ubiquitous communication
- Web technologies and Web caching
- Voice over IP
- Systems and protocols for video, audio, telephony, and games
- Ad hoc and peer to peer networking
- Web protocols and systems, content distribution networks
- Peer-to-peer networking architectures, overlay-based network services and applications, novel distributed applications and middleware
- Streaming services, network storage services
- Programmable network architectures and infrastructure

6.2 Traffic Control and Engineering

- Network management, traffic engineering, and real-world experience
- Quality of service provisioning
- Service pricing
- Overload control

- QoS in multi-provider networks
- Service-aware admission control
- Traffic conditioning
- Routing and switching: algorithms, protocols, and systems
- Mobility management and signalling
- Network planning and optimization: traffic matrix inference, dimensioning procedures
- Network resource management and sharing, operating system support for networking
- Network fault-tolerance, reliability, survivability, debugging, and troubleshooting.
- Traffic measurement: techniques and tools, traffic trends and patterns, definition of performance metrics, load and performance monitoring
- Experimental and measurement results from operational networks and protocols
- Network security, vulnerabilities, and defences
- Network scalability

6.3 Methods and Tools

- Traffic characterization and modelling: characterization at packet, flow and application levels
- Performance evaluation
- Queueing theory
- Scheduling
- Simulation methodology
- Traffic and performance measurements

7. QoS in ongoing/recently terminated European projects and possible follow ups

7.1 Architectures for QoS/SLA monitoring of IP services and applications

7.1.1 TEQUILA - QoS monitoring for value-added IP services

Service Level Specifications (SLS) and Traffic Engineering (TE) are key aspects for the deployment of value-added IP service offerings over the future Internet networks.
Since these IP services are likely to be provided over the whole Internet, their corresponding QoS will be based upon a set of technical parameters that both customers and services providers will have to agree upon. Such agreements, and especially the negotiations preceding them, will be greatly simplified in the presence of an unambiguous set of (technical) SLS parameters. After signing the agreements and specifying the SLSs, it is further the task of the service provider to meet the customer demands through network management and traffic engineering. The customer expects certain performance from the network, but the operator also attempts to satisfy these expectations in a cost-effective manner.
Therefore also in the future traffic engineering is a basic tool for the operator to accommodate as many as possible of the traffic requests by using optimally the available network resources. Basic impact on this research is given by the European project TEQUILA (IST-1999-11253) - Traffic Engineering for Quality of Service in the Internet, at Large Scale http://www.ist-tequila.org/.
TEQUILA's main objective is to study, specify, implement and validate service definition and Traffic Engineering tools for the Internet. The TEQUILA system was intended to provide both quantitative and qualitative service guarantees through planning, dimensioning and dynamic control of traffic management techniques based on DiffServ.
Especially the following technical areas are addressed by TEQUILA:

- Specification of static and dynamic, intra- and inter-domain SLSs (Service Level Specification).
- Protocols and mechanisms for negotiating, monitoring and enforcing SLSs.
- Intra- and inter-domain traffic engineering schemes to ensure that the network can cope with the contracted SLSs - within domains, and in the Internet at large.

The objective of the project is to study, specify, implement and validate a set of service definition and traffic engineering tools to obtain quantitative end-to-end Quality of Service guarantees through careful planning, dimensioning and dynamic control of scaleable and simple qualitative traffic management techniques within the Internet (i.e., DiffServ).
The next figure gives the high-level functional architecture for providing QoS in IP networks as it has been developed within the TEQUILA project. The architecture includes management, control and data-plane functionality. The QoS architecture shows the basic interactions between the provider and the customer, i.e. service subscription, service invocation and data-transmission. The customer may be a company, another (peer) network provider, an application service provider or a residential user.
The functional architecture describing the QoS functionality of the provider contains 5 sub-systems: Service Management, Traffic Engineering, Policy Management, Monitoring and Data-plane functions.

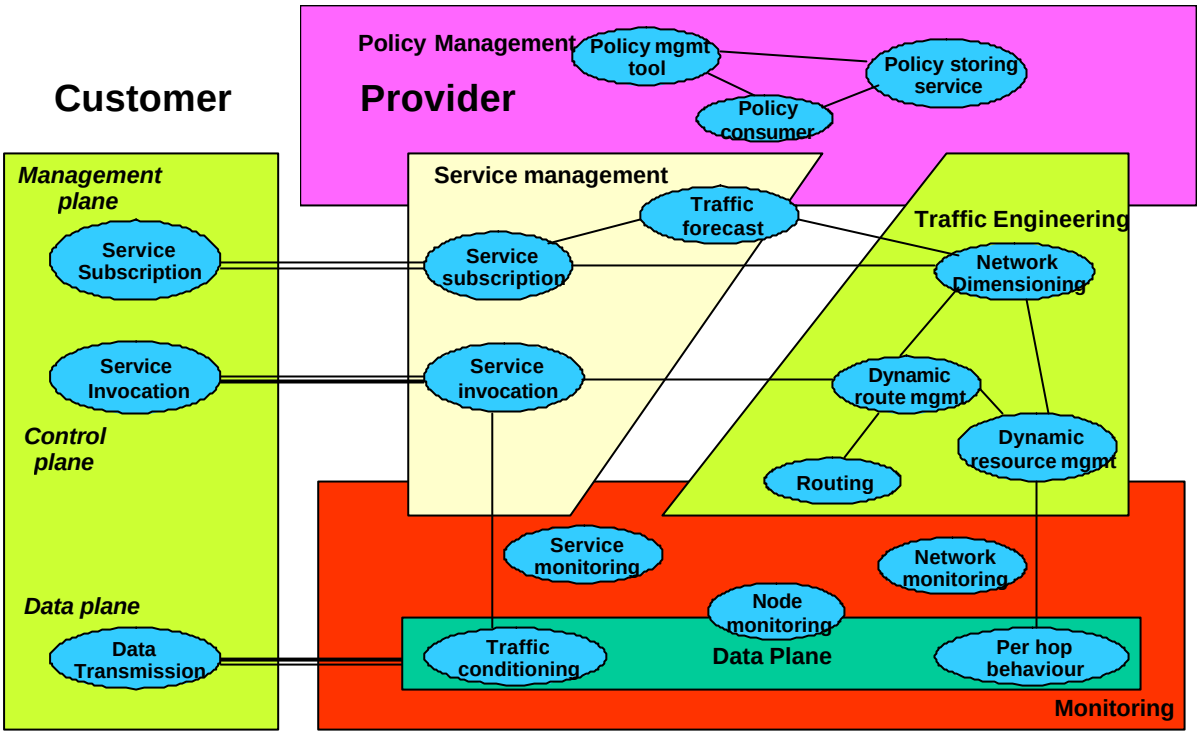

Figure 1 *TEQUILA functional architecture*

The "low-level" **data plane** includes the DiffServ PHB (Per-Hop Behaviour) and TCBs (Traffic Conditioning Blocks), while the "high-level**"** **policy management** allows the administrators to define and enforce policies for both Service Management and Traffic Engineering purposes in an automated way. **Monitoring** is sub-system, which includes node monitoring, network monitoring and service monitoring.

The **Service Management** and the **Traffic Engineering** sub-systems are the essential parts of the system architecture and are the main focus of TEQUILA. Service management includes service creation, negotiation and assurance. Service creation is the process of defining services and service classes by the provider. Service negotiation is the actual negotiation and subscription of value-added IP services between provider and customer. This operational, "on-line" process is the most critical w.r.t. QoS issues, scalability and other resource-related problems, and is one of the main topics addressed by TEQUILA.

Service assurance enables the operator to verify whether the QoS performance guarantees committed in SLAs are in fact being met in its network. This requires an in-service verification of throughput, delay and packet loss characteristics. Service Assurance operates on the statistical data gathered by network monitoring through the network elements.

Traffic Engineering (TE) is the process of specifying the manner in which traffic is treated within the network. TE has both customer and system-oriented objectives. The customers expect certain performance from the network, which in turn should attempt to satisfy these expectations. The expected performance depends on the type of traffic and is specified in the SLSs. The provider on the other hand attempts to satisfy the customer traffic requirements in a cost-effective manner. Hence, the target is to accommodate as many as possible of the QoS requests (as expressed in SLSs) by optimally using the available network resources. This (SLS) service-driven resource management and traffic engineering is another basic TERQUILA research topic. Within TEQUILA, both IP-based and MPLS-based TE techniques are studied.

The TEQUILA architecture emphasises the importance of the Management plane in providing QoS and gives a functional decomposition of the main service and resource management aspects. The key concepts are the following:

- The architecture introduces a two-level approach for (operational) service management and negotiation, i.e. service subscription and service invocation. Both processes occur at a different time scale. Subscription handles the longer term-based service requests that may apply to IP services like IP VPNs, while service invocation acts on a per-call basis, within the context of the deployment of VoIP (Voice over IP) services, for example. The two-level approach in service management is mirrored in the resource management system. The architecture combines a longer-term off-line traffic engineering approach (*network dimensioning component*) with a dynamic on-line handling of traffic fluctuations (the *dynamic route management and dynamic resource management components*).
- The architecture makes a clear distinction between the customer (SLS) aware components and the resource (QoS class) aware components. The Service Management sub-system has the knowledge about the customers, while the Resource Management sub-system knows about the network resources, and acts on the processing of (aggregate) traffics that will be handled by a collection of QoS classes. The inter-working between the two aforementioned sub-systems is clearly defined through the resource provisioning cycle, controlling the interactions between three elementary components of the TEQUILA system: service subscription, traffic forecast and network dimensioning.

The main overall result is that this architecture enables the (dynamic) provisioning of hard QoS guarantees to individual (multimedia) flows while still maintaining a scalable solution. It solves the scalability problem for IP backbones by enabling a two-level approach for admission control.

7.1.2 CADENUS - Creation and Deployment of End-User Services in Premium IPNetworks

The CADENUS project (http://www.cadenus.org/) is aimed to propose an integrated solution for the creation, configuration and provisioning of end-user services with QoS guarantees in Premium IP networks. The solution is based on the CADENUS framework, which is a structuring set of core functional blocks at the user - provider interface. It will provide service creation and configuration in a dynamic way through the appropriate linking of user related service components (authorisation, registration, etc.) to network related service components (QoS control, accounting, etc.). For the provisioning of end-user services with QoS guarantees, a number of components are required which are developed in the project. The project produces recommendations, architectures, mechanisms and policies concerning service configuration and provisioning for both network operators and service providers

- A service configuration and provisioning framework for services with QoS guarantees, which will be developed, implemented, tested and validated in the project.
- Recommendations to network providers and service providers on service provisioning in Premium IP networks which include a mix of network technologies like DiffServ/IntServ, MPLS, ATM, IP/ATM etc.
- A standard way to create and manage Service Level Agreements.

CADENUS is building an integrated solution for the dynamic creation, configuration and provisioning of end-user services with QoS guarantees in Premium IP networks.

Much emphasis is placed on the business processes involved throughout the chain of events, and in this respect the software implementation is based on the commonly used ebXML.
The ebXML framework aims at creating a single global electronic marketplace where enterprises of any size and in any geographical location can meet and conduct business with each other through the exchange of XML based messages. In order for enterprises to conduct electronic business, they must first discover each other and the products and services they have to offer. They then must determine which business processes and documents are necessary to obtain those products and services. Afterwards, they need to find out how the exchange of information will take place and then agree on contractual terms and conditions. Once all of this is accomplished, they can finally exchange information according to these agreements.
The specification of a business process is the main activity required when creating a new service. Afterwards, in order to enable effective negotiation, it is needed that any interested party defines and publishes a Collaboration Protocol Profile (CPP), where a reference to the business process is made, together with the definition of the role that the party wants to play inside such a process. The CPPs, in turn, form the basis for Collaboration Protocol Agreements (CPAs) established between business parties. Ultimately, the business processes specified in the CPAs drive the business service interfaces to execute those processes and send the required documents.
Network Management aspects (especially the relationship with the TMForum's Telecommunication Operators Map) and security (at all levels throughout the architecture) are also addressed.
The CADENUS solution is based on an architecture, which includes key functional blocks at the user-provider interface, within the service provider domain, and between the service provider and the network provider. The capabilities of these functional blocks are reflected in the corresponding SLAs/SLSs.
The three key components in this process are: Access Mediator, Service Mediator and Resource Mediator. The overall mediation procedure includes the mapping of user-requested QoS to the appropriate service-/network- resources, taking into account existing business processes. Contributions in this area, and on the corresponding SLAs and SLSs are being made to the IETF.

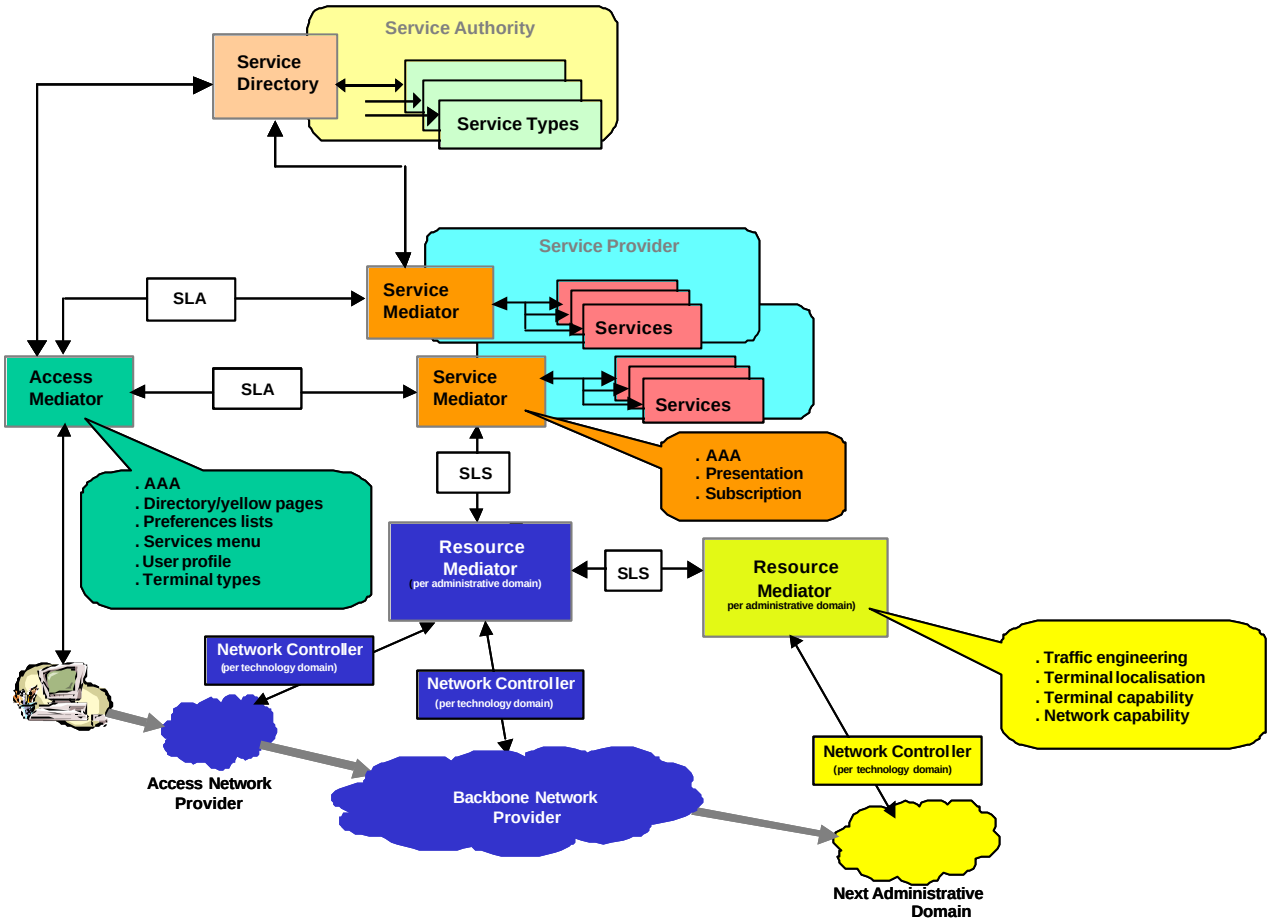

Figure 2 –CADENUS architecture

The key contributions of CADENUS are:

1. A**rchitecture** in which the relationship can be seen between end-user services requiring QoS, and the Premium IP network transport services used to deliver these services. Resources reserved on registration/subscription, and those that are used - and subsequently modified - when the service is invoked/configured are taken into account.
2. Definition of c**omponents** of the architecture such as Access Mediator, Service Mediator and Resource Mediator.

CADENUS uses both MPLS and DiffServ networks as appropriate examples on which to validate their service selection, configuration and creation architecture. This architecture relies on interactions with the management interfaces of such networks for the reservation of resources that guarantee the QoS for the various services.

7.1.3 AQUILA - QoS architecture for adaptive resource control

Adaptive Resource Control For QOS in AQUILA www.ist-aquila.org is aimed to enable **dynamic end-to-end QoS** provisioning in IP networks for QoS sensitive applications e.g. Internet telephony, premium web surfing and video streaming. Static resource assignments will be considered as well as dynamic resource control.
The project assumes the DiffServ architecture as the most promising starting point for its work. The project develops extensions of this architecture in order to avoid the statically fixed pre-allocation of resources to users. Dynamic adaptation of resource allocation to user

requests is enabled in a way that keeps the overall architecture scalable to very large networks.

The Resource Control Layer (RCL) is an overlay network on top of the DiffServ core network The Resource Control Layer provides an abstraction of the underlying layers. The RCL mainly has three tasks, which are assigned to different logical entities:

- to monitor, control and distribute the resources in the network by the Resource Control Agent (RCA).
- to control access to the network by performing policy control and admission control by the Admission Control Agent (ACA).
- to offer an interface of this QoS infrastructure to applications by the End-user Application Toolkit (EAT).

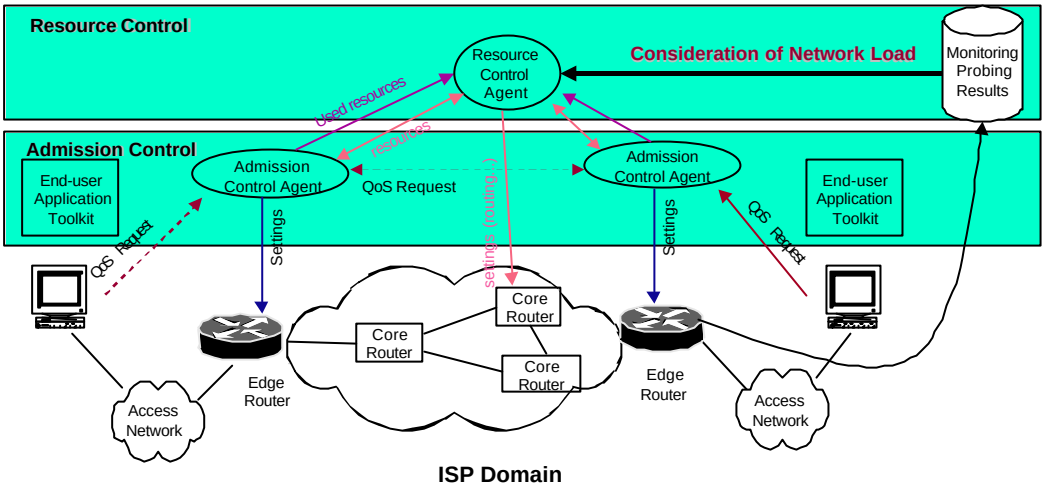

Figure 3 - *AQUILA Architecture*

A node in the Resource Control Layer is called a Resource Control Agent and represents a portion of the IP network, which internally has the same QoS control mechanisms. An RCA is a generalisation of the concept of the Bandwidth Broker in the DiffServ architecture. RCAs are logical units that run on several physical configurations, e.g. one server per RCA or several RCAs co-located on one server. The QoS control mechanisms used in the underlying network are of varying nature, e.g. in some part the routers may not even support DiffServ - which means that there is only a trivial best-effort QoS control - while in other parts they may be DiffServ capable. Moreover, some parts of the network may allow dynamic reconfiguration of resources, e.g. by adding ATM connections, others may have a more or less fixed configuration, e.g. pure SDH or WDM sub-networks. Another reason for the introduction of separate RCAs is that sub-networks are domains managed by different operators.

A Resource Control Agent is able to observe and in some sense to influence the actual configuration in the network portion it represents. Configuration parameters may describe the fraction of a network connection devoted to a specific DiffServ traffic class or the existence of a virtual connection (in ATM networks) with a specified bandwidth.

A DiffServ network can only provide Quality of Service, if it is accompanied by an admission control, which limits the amount of traffic in each DiffServ class. The AQUILA architecture uses a local admission control located in the Admission Control Agent, which is associated with the ingress and egress edge router or border router. To enable the ACA to answer the admission control question without interaction with a central instance, the RCA will locate

objects representing some share of the network resources nearby the ACA. Resources are assigned to these objects proactively.
Admission control can be performed either at the ingress or at the egress or at both, depending on the reservation style.
The ACA will just allocate and de-allocate resources from its associated share. The ACA is not involved in the mechanisms used by the RCA to provide this resource share, to extend and to reduce it.
Resources are handled separately for incoming traffic (ingress) and for outgoing traffic (egress). The following description of resource distribution applies to both.
Resource distribution is performed by the RCA in a hierarchical manner using so-called *Resource Pools*. For this purpose it is assumed, that the DiffServ domain is structured into a backbone network, which interconnects several sub-areas. Each sub-area injects traffic only at a few points into the backbone network. This structuring may be repeated on several levels of hierarchy.
The End-user Application Toolkit (EAT) aims to provide access to end-user applications to QoS features. The EAT is a middleware between the end-user applications (Basic Internet Applications and Complex Internet Services) and the AQUILA network infrastructure.
The EAT supports two major kinds of (Internet) applications:

- Legacy Applications that are in fact QoS-unaware and that cannot be modified in order to directly access the EAT or any other QoS infrastructure. The most of existing Internet applications are legacy ones.
- QoS-aware Applications that can themselves request for QoS, by using an API, for example (EAT-based Applications use the EAT API), or by using signalling protocols such as RSVP and SIP.

Internet applications, however, have also to be distinguished with regard to their *complexity*. In AQUILA, we make a distinction between Basic Internet Applications and Complex Internet Services. They have to be supported in different ways: Whereas Basic Internet Applications are often legacy ones which cannot directly use the EAT, Complex Internet Services can be QoS-aware or even EAT-based although they consist of basic applications.
Generally, the EAT provides – at the control plane – a set of application interfaces in order to support the wide range of different applications:

- ***Legacy applications*** do not interact with the EAT. QoS reservations must therefore be made manually. For that reason, the EAT offers some **Graphical User Interfaces (GUIs)** for manual reservation requests (see below).
- For some ***specific legacy applications*** that dynamically negotiate data port numbers or rely on signalling protocols, special **Protocol Gateways (Proxies)** (e.g. for H.323, SIP) enable the selective processing of the application's control plane information by forwarding QoS-relevant data to the EAT Manager in order to initiate QoS requests. The Proxy Framework is flexible and extensible in order to include additional Proxies (e.g. for RSVP) later on.
- For ***QoS-aware, EAT-based applications***, an **Application Programming Interface (API)** provides interfaces and methods for login, reservation requests and releases, etc. This proprietary API is accessible via CORBA and provides the full AQUILA functionality. The **EAT Manager** directly implements the API in order to manage user access and reservations. (The EAT Manager is the main part of the EAT and controls the whole process. It also acts as mediator between the other EAT components and towards the ACA.)

Due to the fact that the EAT is fully transparent for legacy applications – even if they are supported by a Proxy – QoS reservations must be performed in a different way. For that reason, the EAT provides a set of GUIs in form of Web pages (the so-called **AQUILA Portal**), in which an end-user can *manually* request for QoS reservations. Moreover, the so-called AQUILA Portal offers among other things two different reservation modes: an advanced one for end-users that have knowledge about the technical details of an AQUILA request, and a regular one for end-users that are not familiar with AQUILA.
In order to support the regular reservation mode, an additional application "interface" is provided, the so-called **Application Profile** methodology. Application Profiles contain reservation "schemes" with technical parameters mapped to well understandable QoS metaphors. The **Converter** is the component that takes care of the mapping/converting of the technical parameters of the profiles and the (by the end-user subscribed) network services into the QoS metaphors corresponding to the application in use.
Note that the regular reservation mode is not necessarily part of the AQUILA Portal. In fact, Application Profiles are usable via the EAT API and can therefore be called by every Complex Internet Service that wants to make use of the AQUILA QoS capabilities. In that way, such an Internet service may offer its own regular reservation mode, by showing the QoS metaphors from the proper Application Profiles of its basic applications/plug-ins.
The following figure gives an overview on the above mentioned interfaces and components of the EAT, and how they interact:

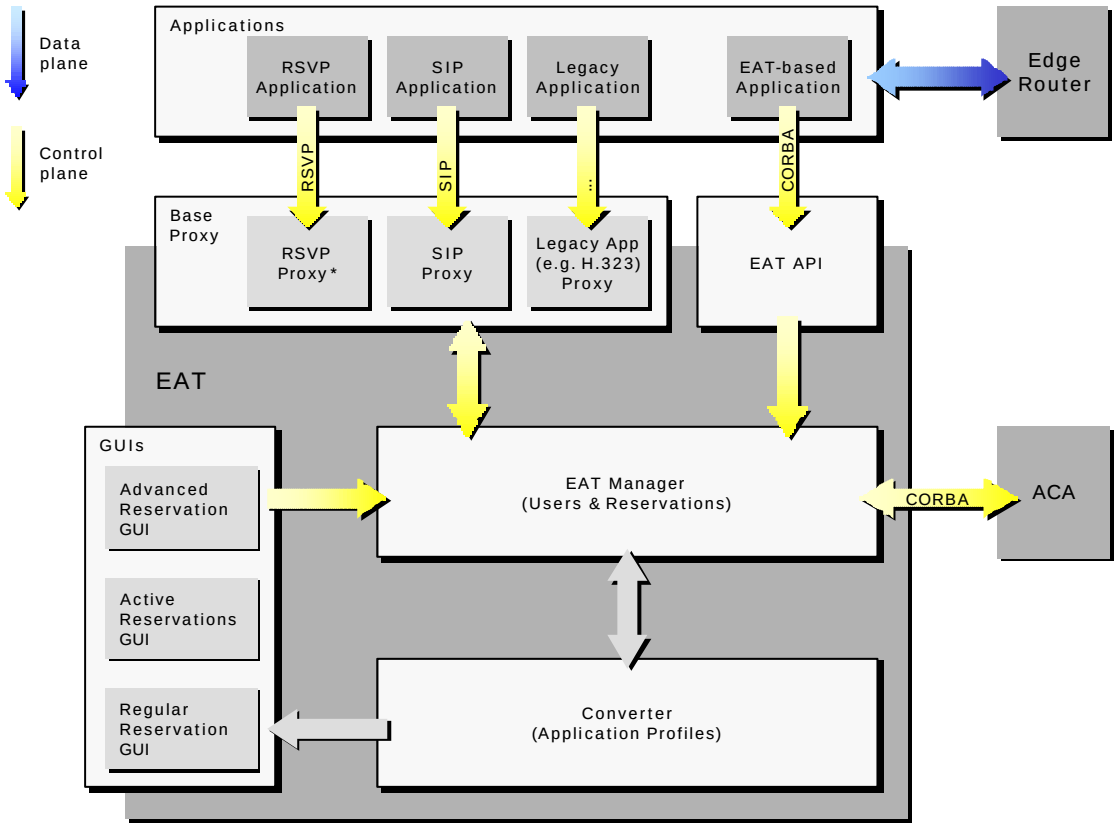

Figure 4 *EAT's basic building blocks and application interfaces*

7.1.4 QUASAR - QoS architectures for core/backbone networks

Quasar project is aimed to define QoS architecture for G-WiN, the German scientific community's Gigabit network The Quasar project proposes a QoS architecture that is qualified for the use in core/backbone networks, in access networks, and in customer premises networks by the definition of well co-ordinating components. There is strong emphasis on inter-operability with other network providers.
The comparison and evaluation of the different QoS approaches for IP networks is the central task of the Quasar project. However, these approaches only make sense if they are supported by a concept for tarifing and pricing. Therefore, Quasar is complemented by concepts for the measurement of the resource usage and a model for authentication, autorisation and accounting (AAA).

7.2 Integrated QoS research

7.2.1 INTERMON inter-domain integrated QoS modelling and visual data mining

The INTERMON architecture www.ist-intermon.org is aimed at integrated QoS (Quality of Service) monitoring, analysis and modelling of application traffic in inter-domain environment using data base and visual data mining facilities. INTERMON features address the automated measurement and modelling of QoS and border router traffic for different time scales as well as visual data mining relating statistics and models describing end-to-end and inter-domain QoS as well as border router traffic. Concepts like ”spatial composition” of inter-domain QoS and “policy-based” performance measurement and traffic collection at border routers are considered. The key points of INTERMON architecture are focussed on:

- Integration of tools for automated Internet structure analysis, monitoring, modelling and visual data mining using common data base for the purpose of QoS monitoring and verification in an inter-domain environment
- Data base design and data mining to support requirements for spatial composition of inter-domain QoS and automated producing of monitoring, modelling and analysis reports considering different aggregation intervals
- Open architecture concept with flexible import/export interfaces for measurement and modelling data (QoS, traffic).

Tools for distributed measurement, modelling and visual data mining using relational data base are integrated in the INTERMON toolkit. INTERMON architecture is intended to be used by ISP provider, operator and application user in an inter-domain environment especially based on QoS technologies (DiffServ, MPLS) for:

- inter-domain traffic engineering and network planning based on visual data mining of border router traffic flows obtained by IPFIX interface
- QoS/SLA monitoring and verification of applications in inter-domain environment based on the concept of spatial composition of inter-domain to end-to-end QoS.

The functional components of the INTERMON toolkit are integrated based on common data base relating topological, measurement and modelling information for different kind of parameters (end-to-end QoS, inter-domain performance metrics, traffic) and Graphical User Interface (GUI).
The integrated measurement and modelling concept is based on relating of the modelling entities (derived per end-to-end QoS parameter, inter-domain performance metric and border router traffic flow) to the corresponding measurement statistics and result entities (describing aggregated statistics results) of the specific measurement scenario for a given time aggregate. The modelling entities, such as accumulative distribution, autocorrelation function, and Auto Regressive Integrated Moving Average (ARIMA) prediction models, are linked to the measurement results and statistics for a given measurement aggregation interval.

INTERMON data mining functions are specified in order to obtain automated generation of modelling reports for a different kind of aggregation intervals (e.g. short term : inter-domain routing, long-term : network planning).
The policy based data collection in INTERMON is open for interoperation with other QoS architectures and flexible to integrate different kind of measurements and statistics into the relational data base.
INTERMON open architecture design is intended to support import/ export measurement and modelling interfaces between different INTERMON users and towards other QoS monitoring and modelling systems.
The INTERMON monitoring tools use remote meters and adapters for execution of measurement/monitoring scenarios specified with the monitoring tools. The remote meters are configured by the monitoring tools to filter their results and to parameterise adapters for the specified measurement scenarios in order to interact with the INTERMON data base. The adapter concept allows reusing the great amount of QoS monitoring data obtained by other QoS monitoring architectures developed in European projects and international activities.

7.2.2 SEQUIN - SErvice QUality across Independently managed Networks

SEQUIN project http://www.dante.net/sequin/ is giving impulse for operational concepts on "Service Quality across Independently Managed Networks". ,
The objective of SEQUIN is to define and implement an end-to-end approach to Quality of Service (QoS) that will operate across multiple management domains and will exploit a combination of IP and ATM technology.
SEQUIN will ensure that researchers across Europe have access to networking facilities that can be tailored to the requirements of the individual groups, and which will offer predictable and stable quality across multiple underlying management domains and networking technologies.
SEQUIN goal is the creation of a definition of QoS which is based on a merging of user requirements and the capabilities of emerging technologies. Especially the evolution of mechanisms for providing Quality-of-Service (QoS) over the contemporary network infrastructures has and the need for regulation and management of the emerging QoS services with the use of Service Level Agreements (SLAs) are addressed by SEQUIN.
SLAs for QoS-enabled networks move one step forward in the direction of traditional ones, in the sense that they do not only have to specify availability, security, quantity of allocated resources and a number of other quantitative values but also have to specify the values of appropriate quality parameters. SEQUIN project is contributing to SLAs used to provide a QoS service called 'IP Premium' from a backbone network (GEANT- the Next Generation of pan-European Research Network to all its peering domains (the National Research Networks-NRENs). The IP Premium service itself was based on the Expedited Forwarding Per Hop Behavior (EF-PHB) of the DiffServ architecture and was defined in the framework of the GEANT and SEQUIN IST projects. The implementation architecture for the Premium IP service aims at offering the equivalent of an end-to-end Virtual Leased Line (VLL) service at the IP layer across multiple domains. SLA specification for QoS enabled networks aims at providing positive quality guarantees and setting out the limits of the services provided. In networks where QoS is inherently supported (such as ATM) the provision of SLAs comes as a natural delimitation of the relevant parameters.
However, in IP networks where best-effort traffic has no quality guarantees, the introduction of QoS and associated services requires a thorough and accurate engineering of QoS metrics in the SLA specification on top of the guarantees for availability and characteristics of the transport medium, security, fault handling etc.
The analytical computation of such metrics is extremely complex taking into consideration the extensive level of aggregation and more generally the nature of traffic flowing in large interconnection domains. Usually only upper bounds for the relevant parameters can be

defined. Therefore, SLA specification for QoS enabled networks becomes a process where intensive testing and probing of the available infrastructure has to take place, before being able to quantify the QoS offering and include concrete parameters and values in the agreement.

7.3 Ipv6 QoS related issues

GÉANT, 6WINIT, LONG, MIND, 6NET and Euro6IX are project contributing to IPV6 QoS introduction and roadmap in Europe.

7.3.1 6NET - Large-Scale International IPv6 Pilot Network

6NET http://www.6net.org/ is a three-year European project to demonstrate that continued growth of the Internet can be met using new IPv6 technology. It also aims to help European research and industry play a leading role in defining and developing the next generation of networking technologies.
The special focus of 6NET is aimed at:

- Install and operate an international pilot IPv6 network with both static and mobile components in order to gain a better understanding of IPv6 deployment issues. This network will primarily use native IPv6 links (initially running at 155 Mbps and increasing to 2.5 Gbps in the second year), although encapsulation over IPv4 infrastructure may be necessary in some cases.

- Test the migration strategies for integrating IPv6 networks with existing IPv4 infrastructure. These include autoconfiguration, handoff, multihoming, renumbering, virtual private networks (VPNs) and quality-of-service (QoS).

- Introduce and test new IPv6 services and applications, as well as legacy services and applications on IPv6 infrastructure. IPv6 Middleware and User Application Trials include real-time videoconference and media streaming applications, online gaming, relational databases, transaction processing systems, and portal services.

- Evaluate address allocation, routing and DNS operation for IPv6 networks.

- Promote IPv6 technology.

7.3.2 Euro6IX - European IPv6 Internet Exchanges Backbone

Euro6IX http://www.euro6ix.org/ is the larger research project up to now funded by the European IST Program (IST-2001-32161). The goal of the Euro6IX project is to support the rapid introduction of IPv6 in Europe.
Euro6IX project is aimed to research an appropriate architecture to design and deploy the first Pan-European non-commercial IPv6 Internet Exchange (IX) Network. It will connect several regional neutral IPv6 Internet Exchange points across Europe, and achieve the same level of robustness and service quality as currently offered by IPv4 Internet Exchange Networks.
The deployed IPv6 IX infrastructure could be used to research, test and validate IPv6-based applications and services, such as:

- Investigations on the maturity of advanced IPv6 network services, as well as the feasibility of their inclusion in the Euro6IX test-bed, for example CoS/QoS, Mobility, Anycast and multicast, security, multihoming, renumbering, and policy languages.

- The development, porting, adaptation, or enhancement of IPv6 enabled applications, which will be made available for project trials and to third parties.

- The research of the legal implications of the project related to users, networks, and service providers addressing, personal data protection, and privacy concerns about IPv6 addressing.

The network built within the Euro6IX project will be open to specific user groups (existing and to be created), who will be connecting to the Euro6IX network by means of a variety of access technologies – mobile, xDSL, cable – and internetworking with legacy IPv4 networks and services, to test the performance of future IPv6 networks, and non-commercial native IPv6 advanced services and applications. The network's Acceptable Use Policy (AUP) excludes the possibility of carrying commercial traffic.

7.4 QoS for multimedia networking

7.4.1 Pro-Net - QoS for real time audio and video

Pro-Net (Production of Broadcast Content in an Object-Oriented IP-based Network) is a British LINK project which started in January 2000. It is concerned with broadcast production networks which transport real-time audio and video. Pro-Net investigates the possibility of using general purpose networking technologies to realise such networks. The emerging IP Differentiated Services (DiffServ) framework is considered as the basis for Quality of Service support; its suitability for transporting traffic with hard real-time guarantees will be evaluated. Pro-Net considers object-oriented signalling and control mechanisms, based on technologies such as CORBA, Java-RMI while Web-based access to services will be provided.

7.5 QoS for mobile networking

7.5.1 BRAIN - Serving IP Quality of Service with HiperLAN/2

The European project BRAIN *(Broadband Radio Access for IP-based Net-work*s) www.ist-brain.org investigates system concepts to provide wireless broadband access to the QoS-based Internet. As basis for the BRAIN radio access HiperLAN/2 is considered. In this paper we propose functions enhancing the Hiper-LAN/2 radio interface to support IP QoS. The special focus is the system ability to schedule the various connections dedicated for IP traffic according their QoS requirements, whereby the specific characteristics of the wireless access such as transmission error control and link adaptation will be regarded.
An IP *Conver-gence Layer (CL)* tailored to the BRAIN radio access provides the functions needed for mapping the QoS requirements of the individual to the QoS parameters avail-able for their respective *Data Link Control (DLC)* connec-tions. Furthermore, this IP CL has to support segmentation and re-assembly to adapt variable length IP packets to fixed length DLC *Protocol Data Units (PDU*).
Mechanisms like error control by means of an ARQ pro-tocol and dynamic link adaptation aim to reduce the radio specifics on the packet loss rate but introduce additional de-lay and overhead to the radio access system that decreases the capacity.

7.5.2 Moby Dick - Mobility and Differentiated Services in a Future IP Network

In order to evolve 3rd Generation mobile and wireless infrastructure towards the Internet - targeting IST 2000 IV 5.2 "Terrestrial Wireless System and Networks", the project Moby Dick is intended to define, implement, and evaluate an IPv6-based mobility-enabled end-to-

end QoS architecture starting from the current IETF's QoS models, Mobile-IPv6, and AAA framework. A representative set of interactive and distributed multimedia applications serves to derive system requirements for the the Moby Dick architecture in a testbed comprising UMTS, 802.11 Wireless LANs and Ethernet.
Main Objective of Moby Dick regarding QoS are aimed at:

- development of seamless access to existing and emerging IP-based applications.
- architecture design for wireless Internet access by developing new mechanisms for seamless hand-over, QoS support after and during hand-over, AAA, and charging.

For future mobile QoS research MobyDick could contribute in following directions:

- Architecture concepts integrating QoS, IPv6 mobility, and AAA (out of the separate architectural approaches for each component currently provided by the IETF) with respect to wireless issues.
- IPv6-based end-to-end technological approach of QoS to fulfil the requirements of present and future mobile communication services.
- QoS models (e.g. Differentiated Services) research in highly dynamic and heterogeneous network topologies (understanding of QoS models is normally restricted to relatively static environments).
- Charging concept which would enable permanent mobile IP based services on a large scale (a strong requirement related to AAA, but currently not a topic within the IETF).

7.5.3 CORTEX - approach for distributed mobile components

The CORTEX IST project (http://cortex.di.fc.ul.pt/) investigates appropriate architectures and QoS paradigms for the construction of applications composed of collections of what may be called *sentient objects* - mobile intelligent software components that accept input from a variety of different sensors allowing them to sense the environment in which they operate before deciding how to react.

This QoS research takes into consideration that future mission-critical computer systems will be comprised of networked components that will act autonomously in responding to a myriad of inputs to affect and control their surrounding environment. These developments will enable a new generation of applications in areas such as intelligent vehicles, mobile robotics, smart buildings, and traffic management as well as in more traditional areas such as telecommunications management, process control and C^3 (command, control and communications). To accommodate growth and adaptability with respect to number of participants, integration of new services, and QoS issues, to name but a few, new computational models are needed. These models must be more powerful than the client/server model, which does not reflect the autonomy and spontaneity of co-operating entities. Proactive applications need active components, which are able to sense their environment and spontaneously interact and co-operate with others. Moreover, the communication infrastructure supporting these applications will involve a plethora of different network types and media with widely varying attributes concerning addressing schemes, topology, bandwidth and reliability.

A key enabling technology to realise the vision of ubiquitous computing and proactive applications, is an intelligent middleware supporting appropriate computational models for the envisaged generation of applications. Such middleware must support growth and adaptability to new technologies, and has to provide the hooks for these applications to enforce non-functional quality attributes like reliability and timeliness. In particular, the middleware has to cope with applications that have some or all of the following characteristics:

- Sentience – the ability to perceive the state of the surrounding environment, through the fusion and interpretation of information from possibly diverse sensors;

- Autonomy – components of these applications will be capable of acting in a decentralised fashion, based solely on the acquisition of information from the environment and on their own knowledge;

- Large scale - typical applications may be composed of billions of interacting hardware and software components;

- Time criticality - these applications will typically interact with the physical environment, and will have to cope with its pace, regardless of adverse conditions due to scale and technology shortcomings;

- Safety criticality – typical applications will interact with human users, whose well-being will frequently rely on them;

- Geographical dispersion - unlike current embedded systems, typical applications will integrate components that are scattered over buildings, cities, countries, and continents;

- Mobility – furthermore, they must possess the ability to move between hosts possibly of different networks, while remaining in continuous operation

- Evolution – these applications will have to cope with changing conditions during their lifetimes. Not only must the applications be designed to evolve, but their underlying support must also be adaptable.

Traditional approaches to the design of time and safety critical distributed applications cannot handle the complexity inherent in the scale and geographic dispersion of these new applications. On the other hand, new promising approaches, such as *autonomous decentralised systems* - a subject of active research during the past few years are beginning to emerge. The suitability of autonomous decentralised systems is being tested in current attempts to develop applications in areas such as air traffic control, with the free-flight approach, and in the Telecommunications Intelligent Network Architecture (TINA) effort. However, whereas basic technologies exist that make autonomous decentralised systems a possibility, this approach is still far from being mature. Fundamental research still needs to be carried out to address appropriate architectures and paradigms for the construction of these applications.

In the CORTEX approach, applications will be composed of collections of what may be called sentient objects - mobile intelligent software components that accept input from a variety of different sensors allowing them to sense the environment in which they operate before deciding how to react. Sentient objects must be able to discover and interact with each other and with the physical world in ways that demand predictable and sometimes guaranteed quality of service (QoS), encompassing both timeliness and reliability guarantees. Achieving predictability is made difficult by the characteristics of the changing environment in which these objects operate, including an unstable and mobile object population, unpredictable network load, varying connectivity, and the presence of failed system components. Thus, the construction of applications from sentient objects must take account of the fundamental trade-off between the existence of a dynamic environment and the need for predictable operation.

References

[1] Edward Knightly - Slides presented at QoS ICC panel, NY, June 2002

[2] Mike Bukley's Slides summarizing Audience's debate on QoS at the ETSI QoS Workshop, Munich Feb 2002 – Available at www.etsi.org/qosworkshop

[3] T. Schroeder et al – Scalable Web Server Clustering Technologies – IEEE Network May 2000

[4] S. Ranjan et al. "QoS Driven Server Migration for Internet Data Centers" – Proceedings of IWQoS 2002

[5] www.akamai.com

[6] www.exodus.net

[7] www.mirror-image.com

[8] Carlstrom infocom 2002 - http://www.ieee-infocom.org/2002/papers/560.pdf

[9] Chen Infocom 02 - http://www.ieee-infocom.org/2002/papers/374.pdf

[10] Andrews Infocom 2002 - http://www.ieee-infocom.org/2002/papers/678.pdf

[11] RPR alliance white paper – www.rpralliance.org

[12] 3GPP "Universal Mobile Telecommunication System (UMTS); Qos Concept and architecture. TS 23.107, v3.5.0 http:www.3gpp.org, 1999

[13] Y. Guo et al. „Class-Based QoS over Air Interfaces in 4G Mobile networks" – IEEE comm. Magazine March 2002, pp. 132

[14] M. N. Moustafa et al. "QoS enabled Broadband Mobile Access to Wireline Networks" – IEEE Communications magazine, April 2002, pp. 50

[15] A.T. Campbell et al "Design, Implementation and Evaluation of Cellular IP" – IEEE Pers. Communication, vol 15 no 1 Jan 2001, pp 36

[16] R. Ramjee et al. "HAWAII: A Domain Based Approach for Supporting Mobility in WAN Wireless Networks" Proc. IEEE Conference on Network Protocols, 1999, pp 283

[17] Yu Cheng et al. "Diffserv Resource Allocation for Fast Handoff in Wireless Mobile Internet – IEEE Comm. Magazine, May 2002 p. 130

[18] T. Zhang et al. "Local Predictive Resource Reservation for handoff in Multimedia Wireless IP Network" IEEE Journal on Selected Areas in Communications – Vol. 19, no 10, October 2001 – p1931

[19] www.globalipsound.com

[20] draft-andersen-ilbc-01.txt available on www.ietf.org

[21] Slides presented by Alan Duric at the ETSI QoS Workshop, Munich Feb 2002 – Available at www.etsi.org/qosworkshop

[22] Slides presented by JY Le Saount at the ETSI QoS Workshop, Munich Feb 2002 – Available at www.etsi.org/qosworkshop

[23] Slides presented by Xavier Hatrisse at the ETSI QoS Workshop, Munich Feb 2002 – Available at www.etsi.org/qosworkshop

[24] Slides presented by Albert Vermeire at the ETSI QoS Workshop, Munich Feb 2002 – Available at www.etsi.org/qosworkshop

[25] http://www.radware.com/content/products/fire.asp

[26] http://www.opensourcefirewall.com/ddos_whitepaper_copy.html

[27] http://www.cert.org/

[28] Achim Autenrieth and Andreas Kirstädter, "Engineering end-to-end IP Resilience using resilience-differentiated QoS", IEEE Communications Magazine, vol. 40, no. 1, Jan 2002 pp. 50-57

[29] D. O. Awduche, A. Chiu, A. Elwalid, I. Widjaja, X. Xiao, Overview and Principles of Internet Traffic Engineering, draft-ietf-tewg-principles-00.txt, Sept., 2001.

[30] A. Feldmann, A. Greenberg, C. Lund, N. Reingold, and J. Rexford, "NetScope: Traffic engineering for IP networks". IEEE Network Magazine, special issue on Internet Traffic Engineering, March/April 2000, pp. 11-19.

[31] A. Feldmann, A. Greenberg, C. Lund, N.Reingold, J. Rexford, and F. True "Deriving Traffic Demands for Operational IP Networks: Methodology and Experience". IEEE/ACM Transactions on Networking, June 2001.

[32] A. Feldmann and J. Rexford, "IP network configuration for intradomain traffic engineering". IEEE Network Magazine, September/October 2001, pp. 46-57.

[33] X. Xiao, A. Hannan, B. Bailey, L.M. Ni, "Traffic engineering with MPLS in the Internet", IEEE Network Magazine, March/April 2000, pp. 28-33.

[34] P. Aukia, M. Kodialam, P.V.N. Koppol, T.V. Lakshman, H. Sarin, B. Suter, B, "RATES: a server for MPLS traffic engineering", IEEE Network Magazine, March/April 2000, pp. 34-41.

[35] www.ist-intermon.org

[36] IEEE 802.11, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Standard, IEEE, August 1999.

[37] C. Heegard, J. Coffey, S. Gummadi, P. Murphy, R. Provencio, E. J. Rossin, S. Schrum, and M. B. Shoemake, "High Performance Wireless Ethernet", IEEE Communications Magazine, vol. 39, no. 11, November 2001.

[38] T. Nandagopal, S. Lu, and V. Bharghavan, "A Unified Architecture for the Design and Evaluation of Wireless Fair Queuing Algorithms," in Proc. ACM MobiCom'99, Seattle, WA, Aug. 1999.

[39] N. H. Vaidya, P. Bahl, and S. Gupta, "Distributed Fair Scheduling in Wireless LAN," in Proc. ACM MobiCom'00, Boston, MA, Aug. 2000.

[40] J. Khun-Jush, G. Malmgren, P. Schramm, and J. Torsner, "HIPERLAN Type 2 for Broadband Wireless Communication," Ericsson Review, no.2 (see http://www.ericsson.com/review), 2000.

[41] V. Kanodia, C. Li, B. Sadeghi, A. Sabharwal, and E. Knightly, "Distributed Multi-Hop with Delay and Throughput Constraints," in Proc. ACM Mobi- Com'01, Rome,Italy, Jul. 2001.

[42] M. Barry, A. Veres, and A. T. Campbell, "Distributed Control Algorithms for Service Differentiation in Wireless Packet Networks," in Proc. IEEE INFOCOM'01, Anchorage, Alaska, Apr. 2001.

[43] A. Imad and C. Castelluccia, "Differentiation Mechanisms for IEEE 802.11," in Proc. IEEE INFOCOM'01, Anchorage, Alaska, Apr. 2001.

[44] A. Banchs, X. P´ erez, and M. Radimirsch, "Assured and Expedited Forwarding Extension for IEEE 802.11 Wireless LAN," in Proc. IWQoS'02, Miami, FL, Jun. 2002.

[45] IEEE 802.11e/D2.0, Part 11: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), Draft Supplement to IEEE 802.11 Standard – 1999 Edition, 2001.

[46] J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski, "Assured Forwarding PHB Group," RFC 2597, Jun. 1999.

[47] F. Le Faucheur et al., Requirements for support of Diff-Serv-aware MPLS Traffic Engineering, draft-ietf-tewg-diff-te-reqts-01.txt, June, 2001.

[48] D.Mitra and K.G.Ramakrishnan, Techniques for Traffic Engineering of Multiservice, Multipriority Networks, Bell Labs Tech. J. Vol.6, No.1, Jan. 2001, pp 139-151.

[49] J. C. de Oliveira, C. Scoglio, I. F. Akyildiz, and G. Uhl, "A New Preemption Policy for DiffServ-Aware Traffic Engineering to Minimize Rerouting", in Proc. IEEE INFOCOM'02, New York, NY, June 2002.

[50] J. Martin, and A. Nilsson, "On Service Level Agreements for IP Networks", in Proc. IEEE INFOCOM'02, New York, NY, June 2002.

[51] E. Bouillet, D.Mitra and K.G.Ramakrishnan, "The Structure and Management of Service Level Agreements in Networks", IEEE Journal on Selected Areas in Communications, vol. 20, no. 4, May 2002

[52] Ying Xu and Roch Guérin, "Individual QoS versus Aggregate QoS: A Loss Performance Study", in Proc. IEEE INFOCOM'02, New York, NY, June 2002.

[53] http://www.ietf.org/html.charters/ipfix-charter.html

[54] http://www.ietf.org/html.charters/psamp-charter.html

[55] http://imrg.grc.nasa.gov/imrg/

[56] http://www.ietf.org/html.charters/nsis-charter.html

[57] P. Gevros, J.Crowcroft, P. Kirstein, and S. Bhatti, "Congestion Control Mechanisms and the Best Effort Service Model", IEEE Network Magazine, May/June 2001, pp. 16-26.

[58] J. Widmer, R. Denda, and M. Mauve, "A Survey on TCP-Friendly Congestion Control", IEEE Network Magazine, May/June 2001, pp. 28-37.

[59] R. Gibbens, and P. Key, "Distributed Control and Resource Marking Using Best-Effort Routers", IEEE Network Magazine, May/June 2001, pp. 54-59.

[60] P. Hurley, J.-Y. Le Boudec, P. Thiran, M. Kara, "ABE: Providing a Low-Delay Service within Best Effort", IEEE Network Magazine, May/June 2001,pp. 60-69.

Acronyms

AF – Assured Forwarding
BER – Bit Error Rate
CDMA – Code Division Multipe Access
CR-LDP – Constrained Routing Label Distribution Protocol
DNS – Domain Name Service
DSLAM – Digital Subscriber Line Access Multiplexer
EF – Expedited Forwarding
HTTP – Hyper Text Transfer Protocol
IDC – Internet Data Centres
MPLS – Multi Protocol Label Switching
PHB – Per Hop Behaviour
RAN – Radio Access Network
RSVP – ReSerVation Protocol
RSVP – TE – RSVP with Traffic Engineering extensions
SIR – Signal to Interference Ratio
Traffic Engineering – TE
VoIP – Voice over IP
VPN – Virtual Private Network