Virtual Networks in the Internet

J-P. Redlich, M. Suzuki, S. Weinstein {redlich, masa, sbw}@ccrl.nj.nec.com

C&C Research Laboratories, NEC USA, Inc.

4 Independence Way, Princeton, NJ08540, USA

Abstract - **Selective treatments are needed for different types of traffic and different user groups, even in the Internet. Virtual networks provide capabilities for custom routings, priorities, and processing that are not available in either Integrated Services or Differentiated Services. This paper relates virtual networks to public network concepts of traffic grooming, virtual private networks and virtual paths, and shows how these and broader capabilities may be realized in the Internet. It describes a networking strategy incorporating *intelligent routers* that can decouple application QoS needs from core network resource allocations. These routers can be programmed to provide virtual networks, which can satisfy a wide range of application requirements.**

I. INTRODUCTION

A virtual network (VN) is a service concept. It is characterized by a set of communication capabilities for an ensemble of communication calls, flows, or sessions that have something in common with one another. The Internet as a whole may be characterized as a virtual network which provides a best effort data forwarding service, augmented with higher level services, such as name services (DNS), management (ICMP, SNMP) and others, with a protocol stack grounded on IP. Other virtual networks may use the same physical infrastructure, but provide additional services or QoS features. Access to these VNs is usually restricted, as the VN features are not universally available or are reserved for usage by authorized users. Efforts such as COPS [1] and DIAMETER [2] are beginning to define client/server protocols to exchange policy information needed for allocation of resources to virtual networks. This paper provides a tutorial introduction to VN concepts and a new architecture for an IP network edge device, an intelligent router, that can translate implicit or explicit user requests for resources into specific network services.

In general, a virtual network is a customized service entity for an identifiable type of traffic or for an identifiable group of users (VPN). In Section 4 this paper describes a model and software prototype permitting relatively simple creation of virtual networks customized to a very wide range of traffic types, including, for example:

- Traffic destined for a particular host or administrative entity (e.g. a particular Internet Bank Web site).
- Traffic associated with a particular individual or user group (e.g. the CEO of a company).
- Traffic associated with a particular Internet service (SMTP, HTTP, DNS, telnet, ...).
- Traffic associated with a particular application (Internet telephony, MPEG video, credit card transaction, ...).
- Traffic of differing qualities or priorities, such as Internet Differentiated Services

Virtual networks are needed to avoid, on the one hand, the cost and complexity of providing a dedicated physical network for each traffic classification suggested in the above list, and, on the other hand, building a very complex "all services" network. Figure 1 illustrates these three alternatives of overlaid physical networks, overlaid virtual networks, and an all services network (as Broadband ISDN was envisioned to be [3]). Creating VNs is really a modularization strategy for an all services network that is scalable (in services) and not overwhelmingly complex. Some potential benefits, in addition to the architectural benefit described above, are:

- Creation and deletion of special treatments (routes, processing, etc.) for aggregations of traffic, as needed.
- Using different, customized control algorithms in the different VNs.
- Providing built-in separations (enhancing privacy) between traffic aggregations with different sensitivities, purposes, or owners.
- Providing communications services customers with virtual private network (VPN) services in the Internet with the same and greater control capability they enjoy with VPN services in the switched telephone network.

0-7803-5261-0/99/$10.00 © 1999 IEEE

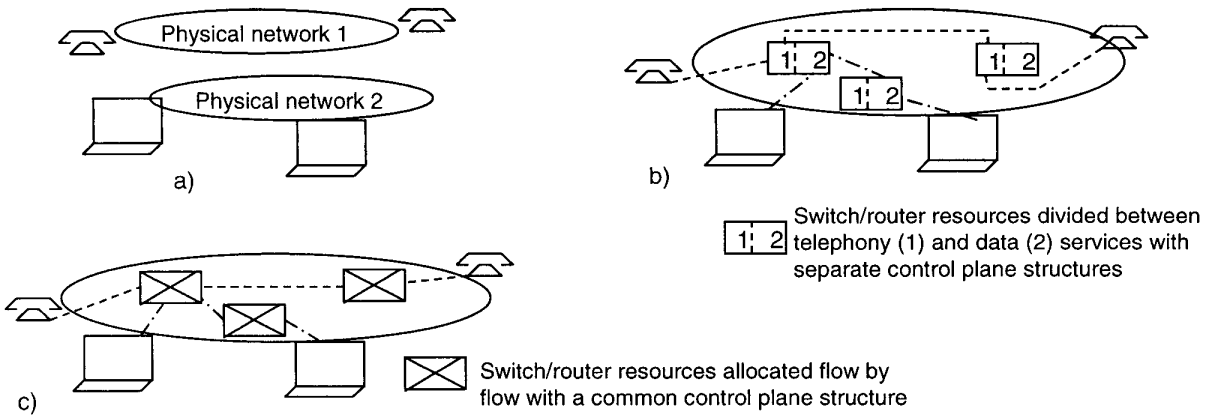

Figure 1: Three ways to realize an all-services communications capability, illustrated for voice and data traffic aggregations. (a) Physical overlay networks. (b) Overlaid Virtual Networks. (c) A single all-services network.

There are, of course, some difficulties and potential disadvantages of VN architectures. First, there is the inefficient use of bandwidth, if pre-allocated in bulk for each VN rather than maintained as one fully shareable resource [4]. Second, it is a severe security challenge to let multiple entities control and manage pieces of the same physical resource [5]. The first concern is being addressed with techniques for quickly reallocating resources among VNs, and the second with operating system-like memory protection and CPU scheduling mechanisms, in conjunction with encryption, authentication, authorization and logging services.

The Internet, in contrast to B-ISDN but, strangely enough, like the telephone network, initially embraced a single-service strategy, best effort IP datagram in the Internet case. The need to offer quality of service (QoS) options has become apparent to the Internet community, but there is still considerable resistance to providing fine-grained per-flow QoS. Virtual networks, which can selectively support (among other possibilities) both per flow Integrated Services and aggregated Differentiated Services [6], are a reasonable framework in which to begin to implement these services and others. Section 4 describes an early study and prototype of this framework.

II. Traffic Customization Precedents : Traffic Grooming, Virtual Private Networks, and Virtual Paths

The concept of virtual networks is already well established in public telecommunications networks, albeit with limited flexibility. Three forms of virtual networks are common:

1) Traffic Grooming
2) Virtual Private Networks
3) Virtual Paths

Traffic grooming is a very old technique used in the telephone network to group traffic with similar characteristics, often related to destination, in order to more efficiently utilize transmission facilities. More recently, SONET add/drop multiplexers have been designed to route local traffic directly from one low-speed port to another, rather than multiplexing and demultiplexing in the high-speed pass-through stream. Voice and data traffic may be separated in digital cross connects (DXCs) and sent through separate trunks to voice and data switches, and Digital Subscriber Line (xDSL) services separate data from voice traffic at the central office. Finally, traffic grooming is appearing in SONET/WDM ring networks as a way of reducing multiplexing costs by grouping similar traffic on particular wavelengths [7].

Virtual private networks (VPNs) are provided to large customers, such as companies linking different locations, who "...do not perceive that they are sharing a network with each other ... you think you have it, but you don't" [3]. A telephone DXC or switch can support the virtual links of several large customers, each of whom may be given some degree of control, particularly in reconfiguring cross connects. The virtual private network is ordinarily changed infrequently and slowly, and is usually restricted to grouping traffic on an owner (source and destination address) basis. Virtual private networks will continue to be an important virtual network category, with more recent work focusing on realizing finer resource dividing strategies and more customized control of routing and other VPN-specific treatments [8].

Virtual paths (VPs) [3] are groupings of ATM virtual circuits (VCs) that are traveling between the same end switches, sometimes groomed so that a VP carries virtual circuits of a particular service class or a particular user or user group. Use of virtual paths allows reuse of virtual circuit identifiers (VCIs) for circuits using the same switches but having different virtual path identifiers (VPIs), and reduces the processing load on intermediate switches which need only process VPIs. Virtual networks are useful in realizing both traffic grooming and VPNs. However, VPs are still just address-controlled routing mechanisms and do not allow the full range of aggregated traffic definitions and treatments that could be possible in virtual networks.

III. Customized Treatment of Traffic in the Internet : Steps toward Virtual Networks

As described in the previous section, the VPN concept of a virtual network, dedicated to a particular group, on top of an already available public network, is not new. It has long been pursued in the Public Switched Telephone Network (PSTN) and VPN services are available today. However, the growth of the Internet has created a new networking environment, which is inexpensive to use and allows access by everyone around the globe. World Wide Web browsers, like the Internet itself, are ubiquitous, making the Internet a common foundation for corporate and private networking.

Current research explores traffic customization in the Internet in two categories. One is service enhancement of the current IPv4-based Internet, and the other is the use of existing services to build a customized networking framework on top of the IPv4-based existing Internet architecture.

The first category includes the Integrated Services architecture [9] and Differentiated Services architecture [10-12]. The Differentiated Services architecture is preferred by many because it scales better in a huge network. Integrated Services, set up using RSVP [13], provide a finer grain QoS guarantee for individual flows, but their feasibility in the global Internet is questioned because states of the flows have to be maintained at each router, which is not necessary with the Differentiated Services architecture.

The Differentiated Services architecture classifies traffic in the network. Every router in the same differentiated services domain has to implement the same Per Hop Behavior (PHB), corresponding to a specific Differentiated Services Code Point (DSCP) embedded in the header of an incoming packet (Figure 2). Six bits are assigned for the DSCP in what used to be the Type of Service (TOS) field [11]. A 32-bit unsigned integer, maintained in the routers, is reserved to define PHB, and a mapping between DSCP and PHB is predefined and maintained at each router. When the router receives an IP packet, it extracts DSCP from the packet's header, looks at the DSCP to PHB mapping table, determines which PHB it should use, and processes the packet according to the predefined PHB.

Differentiated Services can contribute to implementation of virtual networks. The DSCP value in each IP packet assigned to a VN can define the PHB that is part of the special treatment for that VN. This includes access to resources, such as CPU, buffers or space in IO queues. However, to ensure a certain quality of service, the interdependencies of different PHBs which are simultaneously active at the same router, must be limited or should at least be predictable. Future Internet standards will have to address this problem.

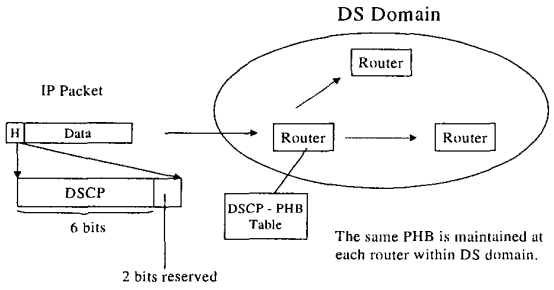

Figure 2: Differentiated Services Architecture.

The second category – building a superstructure for traffic customization - includes the Internet-based VPN (Figure 3). A company may wish to create a VPN on top of the existing Internet at low cost, without fundamentally changing Internet architecture. The VPN may be built from off-the-shelf components such as VPN-enabled routers and firewalls. The main idea of the Internet-based VPN is to have a secured private network on top of the Internet. The key technologies used to realize the Internet–based VPN include:

- Authentication and Encryption.
Those who are involved in a secured communication have to be authenticated at the beginning of their session in order to establish a relation of trust between them. As part of this process, encryption/decryption keys are generated and exchanged. These keys are used by the communication endpoints for transmitting messages over untrusted ground or equipment.

- Tunneling.

Tunneling is a method of using the Internet for connecting two separate (corporate) networks with each other. The IP tunnel is the equivalent to a physical wire. Not only IP packets, but also frames or IPX packets or other layer 2 or layer 3 transmission units can be sent through an IP tunnel. Moreover, since the tunnel provides a direct connection, the two interconnected networks do not have to maintain routing information for the Internet part.

The Security Architecture for the Internet Protocol (IPsec) [14] is a standardized technology that supports both authentication/encryption and tunneling. It enables interoperability among heterogeneous elements in a VPN networking environment. It is applicable to both IPv4 and IPv6. The specification includes two main features, an Authentication Header (AH) and Encapsulation Security Payload (ESP). The AH is inserted between IP header and datagram to provide security parameters (e.g. which authentication algorithm is used). The ESP offers security between two or more hosts implementing ESP by encapsulating the IP header and datagram into an ESP packet.

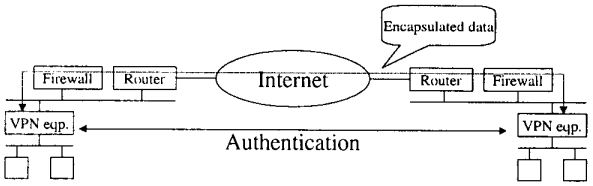

Figure 3 : Internet-based VPN.

A superstructure approach such as Internet VPN can be secure but cannot, unfortunately, achieve QoS goals in a best-effort network. This solution is, therefore, inappropriate for real time traffic such as Internet telephony and conferencing. However, IP tunneling, combined with resource reservation protocols such as RSVP allows for reasonable predictions of the aggregated traffic behavior in the tunnel. Since local area networks at both ends of the tunnel usually over-provision resources, end-to-end QoS issues largely depend on the tunnel and are therefore, within limits, predictable.

Useful traffic routing strategies are possible in the superstructure category. For example, it is possible to set up Internet Service Provider (ISP) access facilities that look like separate networks but actually share the same public network access facilities. One implementation of this idea is the ATM access multiplexer, where configurations of line cards and other subscriber-specific facilities may be allocated to an ISP management entity, but the actual data traffic of the ISP's customers will not necessarily flow through the ISP's own switches and routers [15]. This ISP virtual network helps avoid unnecessary bottlenecks in ISP facilities, letting traffic find its best end-to-end route.

IV. Intelligent Edge Router

One of the Internet's basic design concepts was to push all intelligence into the communication endpoints, i.e. into the end-user's operating system and the applications. The network's sole purpose was to serve as a primarily fast mechanism for exchanging IP packets between these applications. This approach has an inherent advantage over a system where most of the intelligence is provided by the network. It is much easier to provide new mechanisms in new applications than it is to provide new mechanisms inside a network. Application level mechanisms can be implemented spontaneously or by third party software companies . In contrast, changes in network mechanisms have to be planned, standardized, and finally implemented by the network provider's equipment vendors. With most of the intelligence at the network edge, new services with new features can be introduced easily without any need for modifying the underlying network.

However, over time Internet applications have been developed that cannot function properly if the underlying network is unable to commit a minimal set of resources for a specified period of time. These applications, such as Internet Telephony, multimedia document retrieval, or Internet banking/brokering, require a minimum bandwidth, a low maximum transmission delay, or guaranteed reachability (e.g. for a 911 telephone call). New protocols, such as RSVP, allow applications to signal their requirements to network elements along the data transmission path in order to reserve resources there. But new protocols like RSVP are not implemented and widely supported overnight. It may take several years before RSVP is available on a large scale. Today's applications can therefore not use this resource reservation mechanism, even if the underlying networks provide equivalent native mechanisms, for instance the setup of virtual paths with associated QoS.

In addition, an application that signals directly to network elements lacks the global picture. It does not know about other applications or about other users. It does not know which services are most crucial to the end-user, as opposed to those that run just as background entertainment. Moreover the application does not know about the importance of a user within a community such as a corporation. How should an Web-surfer know that he should reduce his resource consumption, because the

marketing department has an IP telephony session with a big potential customer?

To overcome these problems, we propose an architecture that decouples applications from the resource allocation mechanisms of the core network. In this architecture, the resource allocation function is provided by the router that connects a LAN with the Internet core network. Because this router provides programmable functions in addition to its routing and forwarding function, we refer to it as an intelligent router (IR).

The concept of allocating resources to traffic on the basis of attributes of the traffic rather than specific signaling requests is not new. There are very limited realizations in, for example, firewalls that block traffic from certain addresses or applications. What is new here is the flexibility of classification and resource allocation. We can quickly program new criteria for traffic classification, and set up a wide range of treatments for the VNs associated with these traffic classes. Furthermore, we propose an open programming interface for setting up, controlling, and managing VNs from possibly remote locations.

Figure 4 shows our configuration. We assume that for all communication between any two hosts of the LAN, bandwidth and maximum delay meet all application requirements, or if not, that there are mechanisms that can resolve those LAN-resource allocation conflicts. These mechanisms are outside the scope of this paper.

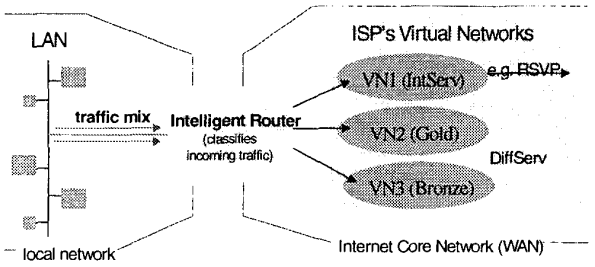

Figure 4: An Intelligent Router (IR) Configuration. (The IR is used to decouple applications on the LAN from the resource allocation mechanisms used inside the core network. A Virtual Network may include processing capabilities in addition to data forwarding service.)

On the other side, the intelligent router is connected to an ISP, which may offer several virtual networks to its customers. Each of these virtual networks may have very specific QoS characteristics, pricing structures and control interfaces. However, as explained above, these virtual networks may share the same physical hardware (e.g. a wire coming out of the wall) or they may actually use a composition of physical access facilities, perhaps belonging to different ISPs. In addition to the virtual networks that are provided by the ISP, the intelligent router may implement its own set of virtual networks, each one with unique, value added functions and interfaces.

The intelligent router is responsible for assigning the resources provided by the ISP to the various flows of IP packets that are emitted from the applications running on the LAN. This resource allocation process is governed by a policy, which is usually defined by the LAN administrator. For a small company's network, this policy may, for instance, require that:

- Traffic originated from the CEO has preference over traffic from staff members.
- FTP traffic during working hours has lowest priority (except for those people that are assigned to a high priority software development project).
- HTTP traffic from summer interns is blocked completely (except traffic to allowed Web sites, such as the company's headquarters web server).
- Email has preference over FTP.

In order to make its decisions, the intelligent router must analyze the traffic it has to forward. The source IP address can be used to determine the user who is associated with this traffic (assuming that an additional component maintains a mapping between users identifications and the IP addresses of their machines). If UDP or TCP is used (which is likely), the port numbers can be used to specify the service/application that is associated with this traffic. For big servers, e.g. of an Internet bank, the destination IP address can be used to determine the associated service. In addition, the application may use signaling, such as RSVP, to indicate its requirements. However, in this case, RSVP is terminated at the intelligent router and is used only to provide additional information about the associated flow of IP packets.

An Internet telephony example helps illustrate how implicit or explicit application requirements are mapped to ISP resources, i.e. to the virtual networks provided either directly by the ISP or by the intelligent router. We assume that the Internet telephony application uses RSVP to signal its requirements for bandwidth and maximum delay to the network. This RSVP signaling is terminated at the intelligent router. Depending on the policies that the LAN administrator defined for the associated user and for the Internet telephony service, the company's CEO may get an ATM switched circuit for his traffic, with bandwidth and delay requirements derived from the RSVP information. Staff members might get the "gold service" from the ISP's

differentiated services virtual network, which in most cases shows good enough behavior for this type of traffic (but which is without any guarantees). Summer interns may run similar applications, but since their traffic is assigned to the "best effort" virtual network they may be not quite satisfied with the QoS they get most of the time. An alternative to using an ATM switched circuit for the CEO's traffic is setting up a path through the Internet core that guarantees a high service quality for the IP telephony session. This assumes, however, that the ISP supports RSVP in its core network.

If the IP telephony application sends its traffic without any additional (RSVP) signaling, the intelligent router applies its default policies without specific knowledge of bandwidth and delay requirements. For instance, for the CEO's telephone call, the intelligent router can use RSVP to reserve nominal resources in the ISP's core network (if RSVP is implemented there), even if the CEO's IP telephony application itself is unable to use RSVP. We see here a compromise between intelligent end systems and intelligent network in order to provide higher quality service even if the end system has not evolved to the latest stage.

Last but not least, the intelligent router could temporarily change its policy in favor of a certain user, if this user either has the privilege to make such changes (service upgrades) or if the user is willing to pay for the higher service level from the ISP. The summer intern from the above example could purchase premium network support for his IP telephony application, even if he would usually have to use the "best effort" network. Assuming the availability of an infrastructure that allows for efficient and secure micro-payments, i.e. transactions below one dollar, the payment for the higher service could either be provided by the user himself or by his remote partner. Hence, as a courtesy to its customers, the summer intern's bank could provide him with high service quality if he accesses his banking account through the bank's Internet site.

The use of the intelligent router for decoupling an application's request for a certain QoS from the network's resource management has the additional advantage that different mechanisms, i.e. protocols, abstractions, etc. can be used in the LAN and in the Internet core network. As alluded in the above examples, new protocols for letting an application express its QoS requirements can be introduced long before such support is available in the Internet core network. Moreover, the intelligent router may use the very efficient native resource management subsystems and signaling protocols of the underlying network in order to meet the applications requirements. Hence, an application may use native ATM without knowing about ATM at all.

V. Implementation Considerations and Conclusions

The implementation architecture is shown in Figure 5. The programming interface defined in OMG IDL includes the following functionalities:

- A programmable pattern factory that produces patterns which are used for classifying traffic on the basis of bit patterns in IP headers.
- Modification of parameters of schedulers for packet forwarding.
- Specification of parameters of token buckets for traffic delimiting.
- Specification of combinations of schedulers and traffic shapers to define VNs.
- Specification of rules for mapping incoming packets into VNs (using patterns).
- Monitoring of packet counts in each operating VN.

The user interface shown in Figure 5 is an application that uses the programming interface to enter or modify assignments of traffic to VNs.

The capacity assigned to each contributing source of traffic for a particular VN is assigned as a fraction of the total throughput measured for that VN. The assignment is implemented with token buckets. The allocations among different VNs can also be modified.

Another component, also shown in Figure 5, is a Domain Resources Manager (DRM) that executes an algorithm for allocating capacity among different VNs. It does this on the basis of traffic congestion information from neighboring routers, obtained through their programming interfaces.

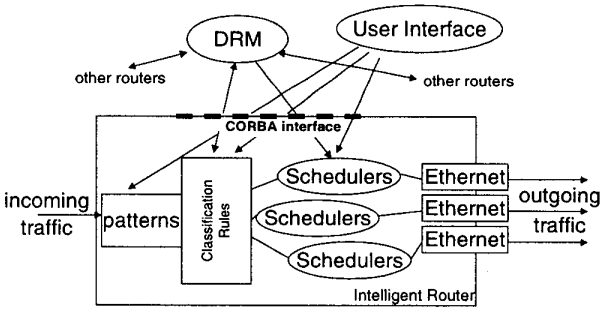

Figure 5: Implementation architecture of a LAN-based intelligent router.

A prototyping testbed is being constructed, using a software router running on a PC with Linux as its operating system. The router components are implemented as CORBA objects, that can be remotely

controlled. With these objects, virtual networks for different classes of traffic have been set up in an Ethernet environment. Initial results show the added burden of using CORBA results in an increased forwarding delay of less than 1ms. This system is intended to be compatible with future CORBA interfaces for network elements following the IEEE P1520 standard [16].

A small group of everyday networking users in our organization are routed through this infrastructure. We can demonstrate fast creation of VNs and changes in assignments of the traffic of different users to these VNs. One application, for example, is on-demand upgrading of a user's VN assignment on the basis of an incremental service charge.

We are developing additional customized traffic treatments in this prototyping environment. We believe that implementations similar to ours will become a common platform for services flexibility in the Internet of the 21st century.

REFERENCES

[1] IETF Internet Draft, "The COPS (Common Open Policy Service) Protocol", Boyle, et. al., November, 1998, available at http://search.ietf.org/internet-drafts/draft-ietf-rap-cops-03.txt.

[2] IETF Internet Draft, "DIAMETER Framework Document", Calhoun, et. al., August, 1998, available at http://search.ietf.org/internet-drafts/draft-calhoun-diameter-framework-01.txt.

[3] U. Black, ATM: Foundation for Broadband Networks, Prentice Hall, 1995, ISBN 0-13-297178-X.

[4] J-F. Huard and A. Lazar, "A programmable transport architecture with QoS guarantees", IEEE Commun. Mag., October, 1998.

[5] S. Alexander, W. Arbaugh, A. Keromytis and J. Smith, "Safety and security of programmable network infrastructures", IEEE Commun. Mag., October, 1998.

[6] P. Ferguson and G. Huston, Quality of Service, Wiley, 1998, ISBN 0-471-24358-2.

[7] E. Modiano and A. Chiu, "Traffic grooming algorithms for minimizing electronic multiplexing costs in unidirectional SONET/WDM ring networks", 1998 Conference on Information Systems Sciences, Princeton.

[8] J. Rooney, J.E. van der Merwe, S.A. Crosby and I.M. Leslie, "The Tempest: A framework for safe, resource-assured, programmable networks", IEEE Communications Magazine, October, 1998.

[9] IETF, RFC1633, "Integrated Services in the Internet Architecture : An Overview", R.Braden, et. al., June, 1994.

[10] IETF Internet Draft, "An architecture for Differentiated Services", S.Blake, et. al., August, 1998, available at http://search.ietf.org/internet-drafts/draft-ietf-diffserv-arch-01.txt.

[11] IETF Internet Draft, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", K.Nicholas, et. al., August, 1998. available at http://search.ietf.org/internet-drafts/draft-ietf-diffserv-header-02.txt.

[12] IETF Internet Draft, "Management of PHBs", M.Borden, et. al., August, 1998, available at http://search.ietf.org/internet-drafts/draft-ietf-diffserv-phb-mgmt-00.txt.

[13] IETF, RFC2205, "Resource ReSerVation Protocol (RSVP) – Version1 Functional Specification", R.Braden, et. al., September, 1997.

[14] IETF, RFC1825, "Security Architecture for the Internet Protocol", R. Atkinson, August, 1995.

[15] R. Dighe, M. Suzuki and S. Weinstein, "The Global Internet: A New Perspective on Broadband Access to the Internet", Proc. IEEE Globecom '98, Sydney, November, 1998.

[16] J. Biswas et.al., "The IEEE P1520 Standards Initiative for Programmable Network Interfaces", IEEE Commun. Mag., October, 1998.