Zeitschrift: Comtec : Informations- und Telekommunikationstechnologie = information and telecommunication technology

Herausgeber: Swisscom

Band: 77 (1999)

Heft: 6

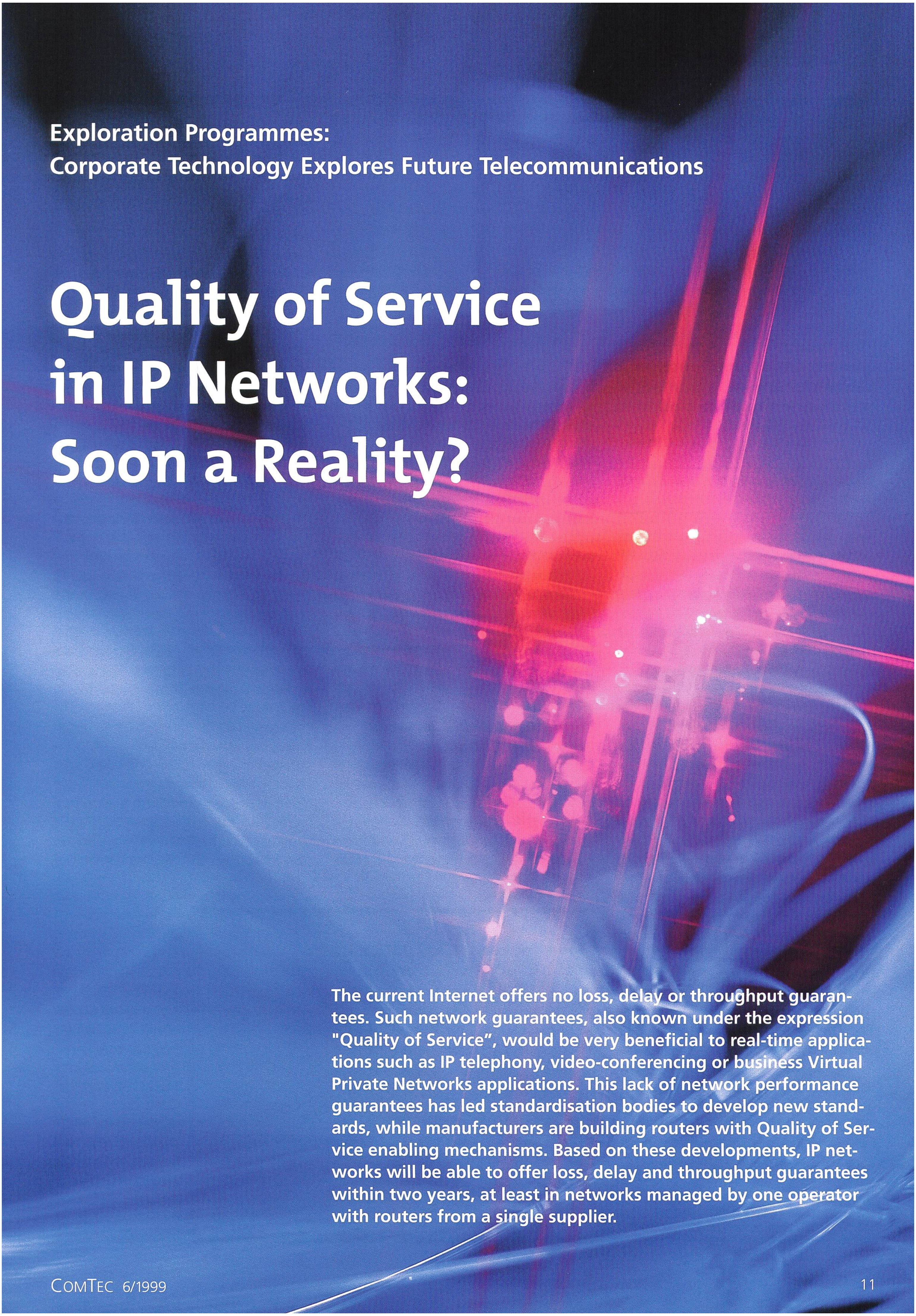
Artikel: Quality of service in IP networks: soon a reality?

Autor: Moix, Dominique / Berger, Fabien / Braun, Fritz

DOI: https://doi.org/10.5169/seals-877027

Exploration Programmes:
Corporate Technology Explores Future Telecommunications

Quality of Service in IP Networks: Soon a Reality?



The current Internet offers no loss, delay or throughput guarantees. Such network guarantees, also known under the expression "Quality of Service", would be very beneficial to real-time applications such as IP telephony, video-conferencing or business Virtual Private Networks applications. This lack of network performance guarantees has led standardisation bodies to develop new standards, while manufacturers are building routers with Quality of Service enabling mechanisms. Based on these developments, IP networks will be able to offer loss, delay and throughput guarantees within two years, at least in networks managed by one operator with routers from a single supplier.

Exploration Programme *"Transport Network Evolution"* elaborates scenarios for optimised use and consolidation of the backbone transport network. The main topic is the economic migration of the network from the voice into the data world. Special emphasis is on the introduction of an optical transport layer and the optimised use of the client layers SDH, ATM and IP. The choice of the needed layers depends on the service portfolio to be offered and has a strong impact on the investment and operation costs of the network, and the flexibility to introduce new services.
With Exploration Programmes Corporate Technology is exploring telecommunication technologies and new service possibilities within projects having a long-term focus of 2–5 years to build up expertise enabling active business innovation support.

Network Performance Guarantees

Network performance guarantees, which can be offered to a customer are of two types: absolute or relative. When absolute guarantees are given, the losses, delay or throughput are bounded and we define these absolute guarantees as Quality of Service (QoS)[1]. ATM networks for instance are capable of providing tight bounds for any of these parameters.

DOMINIQUE MOIX, FABIEN BERGER AND FRITZ BRAUN, BERN

When relative guarantees are given, the traffic is classified into classes and traffic belonging to a class receives a better treatment than traffic belonging to a lower class. We define these relative guarantees as Classes of Service (CoS). The losses, delay or throughput are not bounded. Better treatment can mean for example that the losses experienced by a given class are lower than the losses of a lower class. Other examples of better treatment include lower delays or lower jitter for the higher class. Although the terms QoS and CoS have slightly different meaning, they are often used interchangeably in the literature.
The different performance parameters are related in a complex manner. It is not always possible to offer for example low delays and low losses simultaneously. Indeed, low delays often imply small buffers in the routers. These small buffers have only limited space for data bursts. As a consequence, reducing buffers without changing other parameters generally reduces the delay at the expense of increased losses. Protocols like TCP constantly increase the number of unacknowledged packets sent until a packet drop occurs. TCP needs losses to adapt itself to the available network resources. Thus, loss is not always a good metric for TCP. A good metric for TCP is the achieved throughput.

Programme Scenario

The basic lead question for the Exploration Programme Transport Network Evolution is how to provide a radically more cost-effective backbone network for supporting emerging multi-services markets.

Customer Demand for Network Performance Guarantees

It is not an easy task to evaluate customer demand for network performance guarantees. First, the customer does not always opt for "good" network performance, where "good" network performance is based on a mean opinion score test, for example. The reason is that network performance guarantees are only one aspect of the service offered to the customer. Service pricing and features such as service availability and help-desk assistance also play a very important role in the choice of a particular service provider.
Furthermore, it may be difficult to determine the adequate performance guarantees per application. For data transfer for example, a file is only useful when its transfer is complete. For large files (e.g. some megabytes), a short transfer time may thus require very high throughput. Finally, it should be noted that the offered network performance is not necessarily the one perceived by the end user. During a Web session, for example, the perceived delay includes the time for the DNS look-up and the response time of the server. Typically, the Network Provider has no influence on these delays. In this case, guaranteeing a delay inside the network may be ineffectual.

Emerging Standards

Different standardisation bodies are developing standards and protocols for offering network performance guarantees on the Internet. These standardisation bodies include the Internet Engineering Task Force (IETF) and the International

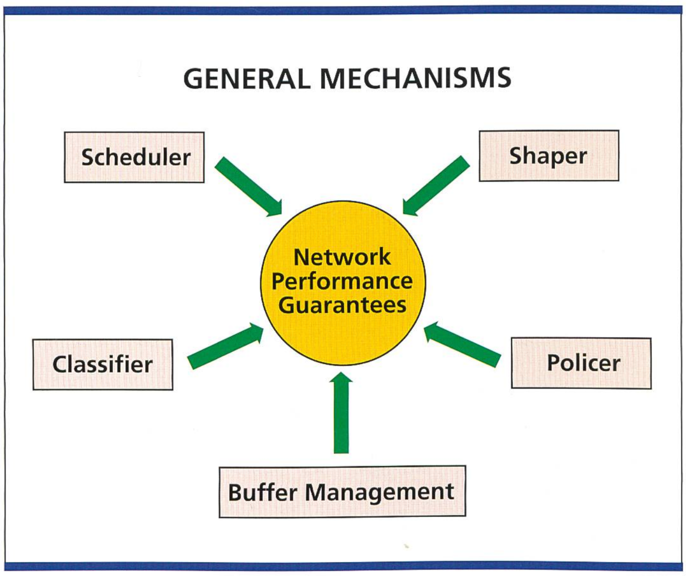


Fig. 1. General mechanisms for offering network performance guarantees: Classifiers, policers, buffer management mechanisms, schedulers and shapers are the building blocks for offering network performance guarantees.

[1] QoS can have different meanings. In this context, QoS refers expecially to the network related parameters loss, delay and throughput.

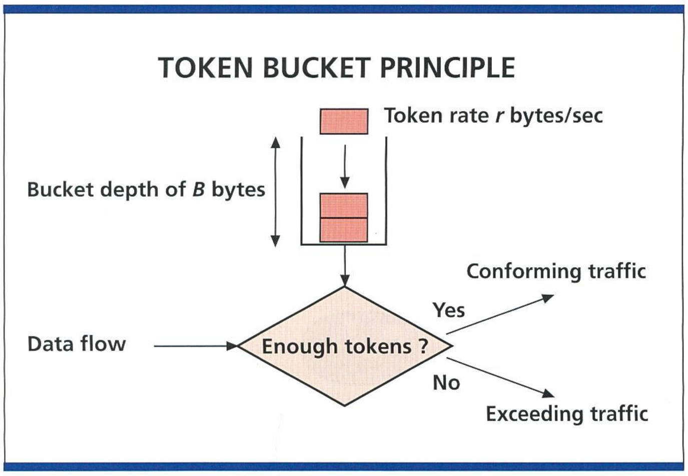


Fig. 2. The principle of the token bucket. An incoming packet is considered as conforming if enough tokens are available in the bucket. If not enough tokens are available in the bucket, the packet is considering as exceeding.

Telecommunication Union (ITU). The IETF is the most active standardisation organisation in the IP area and we will describe its activities in the next sections.

In the IETF, two working groups deal with improving network performance in IP networks. The *Integrated Services* working group pioneered the work in this area by developing a framework, where guarantees are given to individual flows or connections. Two services have been defined: the *Guaranteed Service [1]* and the *Controlled Load Service [2]* The Guaranteed Service provides firm bounds on end-to-end packet queuing delays, whereas the Controlled Load Service approximates the end-to-end behaviour provided by best effort service under low load conditions. The Resource Reservation Protocol (RSVP) [3] is used to reserve resources in the routers. A drawback of this signalling protocol is that it does not scale with large networks: the computational processing and memory consumption in the routers increase in direct proportion to the number of RSVP sessions. As a consequence, the interest for the Integrated Services framework is decreasing for end-to-end flows. However, propositions are being made to use RSVP and Integrated Services in the access whereas the backbone uses a *Differentiated Services* solution.

The most active working group is the *Differentiated Services* group and at this time it receives the largest interest. Its scaleable architecture is a response to the problems of the Integrated Services approach. Here, the traffic is classified into classes and network performance guarantees are given to these classes. The Differentiated Services (DS) field of the IP packets conveys the traffic classification state. The DS field is the new name for the Type of Service (ToS) byte defined in the IPv4 header. The DS byte contains the Per-Hop-Behaviour (PHB) which describes the requested treatment at an IP router. Two Per-Hop-Behaviours are now proposed standards: the *Expedited Forwarding* (EF) PHB [4] and the *Assured Forwarding* (AF) PHB Group [5]. In the EF PHB, the departure rate of the packets must equal or exceed a configurable rate. The EF PHB can be used to build a delivery service characterised by an assured bandwidth with low losses, delay and jitter. The AF PHB defines several classes. The difference between the classes is the delay the packets experience. Within each class, 3-drop probabilities have been defined. In the experiments presented below, we define 2-drop probabilities and show that Service Differentiation can already be achieved.

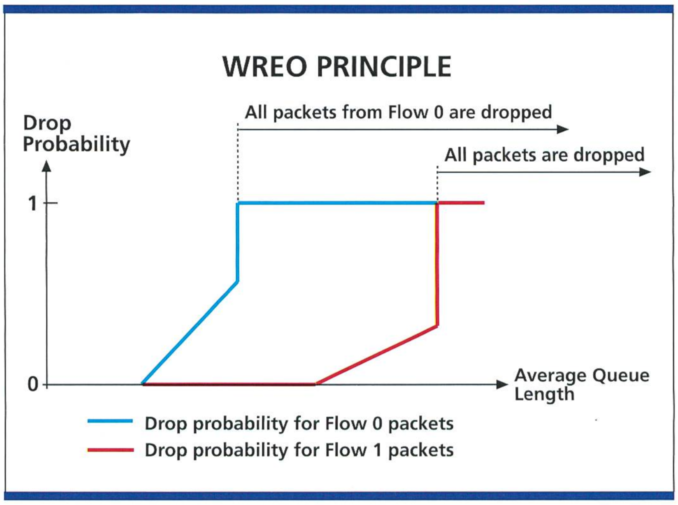
General Mechanisms for Offering Network Performance Guarantees

A router has to implement different mechanisms to offer network performance guarantees. These features include *classifiers, policers, buffer management* mechanisms, *schedulers and shapers*. These mechanisms are the building blocks for the services defined in the *Integrated Services* and *Differentiated Services* working groups (fig.1).

The *classifier* classifies the IP packets into flows, where a flow is based on the source or destination IP addresses, on the protocols used (TCP, UDP), on the port numbers, on the value of the Differentiated Services (DS) field or a combination of these parameters. Thus, a flow can represent a single TCP connection or, at the other extreme, all the traffic between two sites. A flow can also be all the traffic with packets having a given DS field: this is the approach taken by the *Differentiated Services* working group.

The *policer* ensures that only the packets authorised to receive network performance guarantees are preferentially treated and guarantees that misbehaving flows do not degrade the quality of other flows. For doing this, the policer drops or marks exceeding packets belonging to misbehaving flows. Policers are usually implemented with a token bucket. The working principle of a token bucket is depicted in figure 2. The important parameters are the token rate r and the bucket depth B. Tokens are inserted into the bucket at the rate r bytes/s. The bucket depth B is the maximum number of tokens which can be stored in the bucket. An x bytes packet arriving at the token bucket is considered as *conforming* if there are at least x bytes worth of tokens accumulated in the bucket. If not enough tokens are available, the packet is considered as *exceeding*. When an x bytes packet is conforming, x bytes are

Fig. 3. The principle of the weighted random early detection. The drop probability for a packet arriving at a queue increases with the average queue length and depends on the flow. In this example, we have two flows. Flow 0 could e.g. represent exceeding TCP traffic, whereas Flow 1 could represent conforming TCP traffic.



removed from the bucket. The token bucket allows small bursts of data to be considered as conforming.
After the policer, the packets are stored in different queues according to the flows they belong to. It is possible that a given queue is becoming full and packets which should go into this queue have to be dropped. *Buffer* management mechanisms determine which packets are dropped. The simplest method is to drop packets only when the queue is full. This method, called *Tail Drop,* has the disadvantage that losses for TCP connections will occur simultaneously. As a consequence, the TCP connections will also back off simultaneously, thus leading to a lower throughput. A solution to this problem is to increase the drop probability of packets with the average queue size, as in the *Random Early Detection* (RED) mechanism. If multiple flows are directed to a single queue, it is possible to set different drop probabilities for the different flows. This mechanism is called *Weighted Random Early Detection* (WRED) and allows performance differentiation for traffic buffered in a single queue. A WRED configuration for two flows is depicted in figure 3.
The *scheduler* determines in which order the packets are transmitted over the output link. Different mechanisms exist to determine which queue may send: in *Priority Queuing,* all the packets of a higher priority queue are transmitted before packets from a lower priority queue are transmitted. With this approach, packets from a lower priority queue may be delayed indefinitely. Better scheduling mechanisms are *Weighted Fair Queuing* (WFQ) algorithms. With these algorithms, each queue is served at a rate which is at least a given share of the output link bandwidth. If a queue is empty, the other queues may use the bandwidth that the empty queue is not using. WFQ algorithms offer a bounded maximum delay. A drawback of WFQ algorithms is that they often require large computing resources. A WFQ example is shown in figure 4.
Priority Queuing or WFQ determine the *queue* which may send a packet on the output link. They do not determine *which packet* in the queue will be sent. In a given queue, the most common algorithm used for determining the next packet to be sent is *First In First Out* (FIFO): among all packets of the queue, the first arrived packet will be sent.
Shapers ensure that flows become compliant to the Service Level Agreement

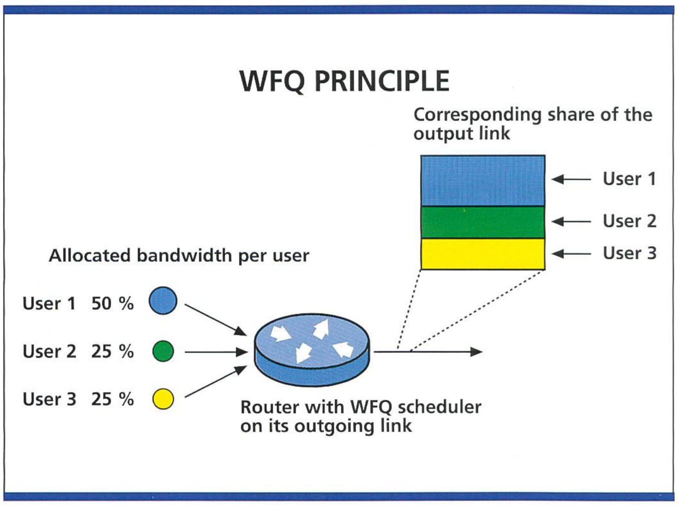


Fig. 4. The principle of weighted fair queuing. The output link is shared among the users according to the allocated bandwidth.

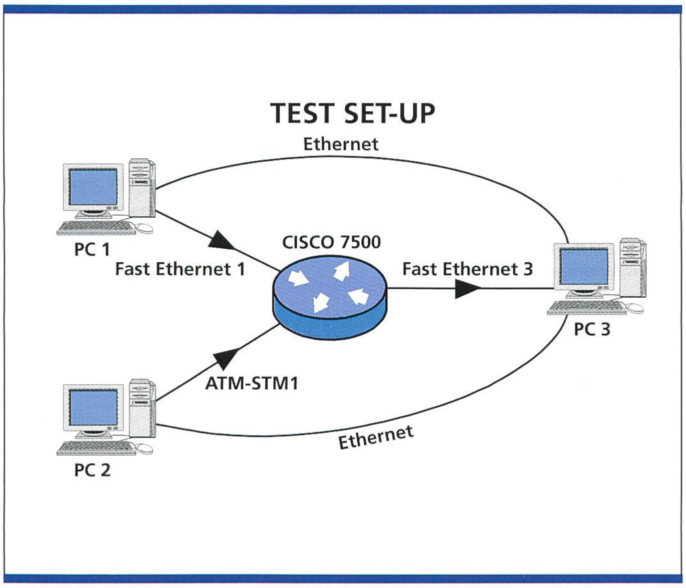


Fig. 5. Test set-up. To investigate service differentiation, 3 PCs and a router were connected together. On the input interfaces of the router (Fast Ethernet 1 and ATM-STM-1), token bucket policers were used to mark the incoming packets as conforming or exceeding. On the congested interface Fast Ethernet 3, Weighted Random Early Detection was applied for obtaining service differentiation.

Abbreviations

AF	Assured Forwarding
ATM	Asynchronous Transfer Mode
CoS	Class of Service
DNS	Domain Name System
DS	Differentiated Services
EF	Expedited Forwarding
FIFO	First In First Out
IETF	Internet Engineering Task Force
IOS	Internetworking Operating System
IP	Internet Protocol
ITU	International Telecommunication Union
PHB	Per-Hop-Behaviour
QoS	Quality of Service
RED	Random Early Detection
RFC	Request For Comments
SLA	Service Level Agreement
TCP	Transmission Control Protocol
ToS	Type of Service
UDP	User Datagram Protocol
VPN	Virtual Private Network
WFQ	Weighted Fair Queuing
WRED	Weighted Random Early Detection

References

[1] Shenker, S., Specification of Guaranteed Quality of Service, RFC 2212, 1997

[2] Wroclawski, J., Specification of the Controlled-Load Network Element Service, RFC 2211, 1997

[3] Braden R. et al., Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification, RFC 2205, 1997

[4] Jacobson, V. et al., An Expedited Forwarding PHB, Internet Draft, February 1999

[5] Heinanen, J. et al., Assured Forwarding PHB Group, Internet Draft, February 1999

(SLA) contracted with the service provider. This is obtained by buffering the packets in a queue and reading them out of the queue at a given rate. When the queue is full, the packets are dropped. With a shaper, packet bursts are smoothed out. The shaper approach gives better performance for TCP connections than a policer does.
The mechanisms described above for offering network performance guarantees are not specific to IP. Indeed, most of these mechanisms are already used in today's ATM networks.
The EF PHB described above can be implemented by combining a token bucket policer with a scheduler based on priority queuing. The AF PHB described above can be implemented by combining a token bucket policer, a WFQ scheduler and WRED.

Example of Service Differentiation

In this section, we show an example of service differentiation as proposed by the *Differentiated Services* framework. This example is one of the results of tests we performed in our laboratories. The goal of these tests was to demonstrate that service differentiation can be achieved by combining the mechanisms described above. We used a Cisco 7500 router running the special CoS release IOS version 11.1.20CC. The router was equipped with 2 Fast Ethernet (100 Mbit/s) and one ATM (155 Mbit/s) interfaces. Each interface was connected to a PC as described in figure 5. During the tests, large files were transferred with TCP between PC 1 and PC 3, and between PC 2 and PC 3. The only limitation was the TCP protocol and the underlying network, not the application above. The Ethernet links were used to send back the TCP acknowledgements from PC 3 to PC 1 and PC 2. Thus, no collisions were occurring on the Fast Ethernet link between the router and PC 3.
Three customers were on PC 1 and three others on PC 2. Each customer was trying to send a large file to PC 3 with the use of TCP. Each customer used one TCP connection. The 6 TCP connections were running simultaneously during 30 seconds. The customers were split into Premium and Base customers. We wanted to achieve service differentiation by giving different throughputs to the different customers. Premium customers got twice more bandwidth than Base customers. For doing this, we used a token bucket policer on the input interfaces of the router (Fast Ethernet 1 and ATM-STM-1): conforming packets were marked with a given DS byte and exceeding packets with another DS byte. We call the token rate of the Base customers the base token rate. For the Premium customers, the token rate was set to twice the base token rate: thus, for example, if the base token rate is set to 1 Mbit/s, Premium customers get 2 Mbit/s. For all customers, the bucket depth was set to 2 Mbytes and was left unchanged during all the measurements. A measurement was done for a given base token rate. Then, the base token rate was changed and the measurements repeated.
On the congested interface Fast Ethernet 3, we used WRED for obtaining service differentiation. Thus, the drop probability for conforming packets was smaller than the drop probability of the exceeding packets. The combination of a token bucket policer with WRED corresponds to a special case of the AF PHB: in our case we have only one class, in opposition to AF PHB which defines 4 classes.
The result of the experiment we ran with 6 customers is depicted in figure 6. The throughput for each customer is displayed versus the base token rate. By varying the base token rate, we vary the network provisioning level: we per-

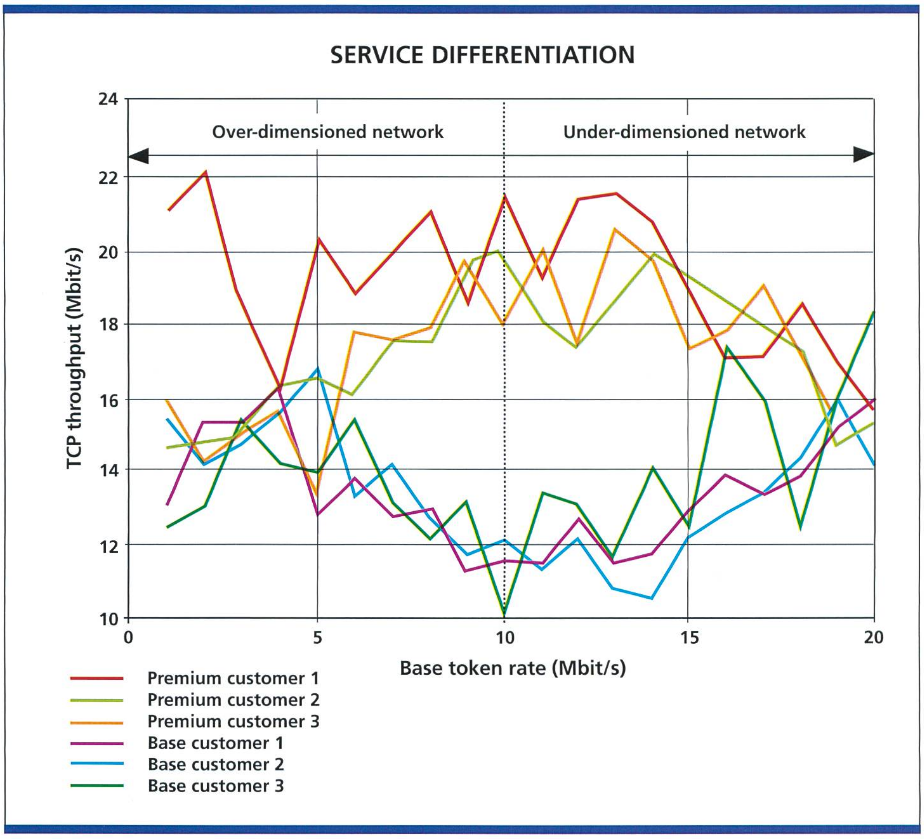


Fig. 6. An example of service differentiation. Premium customers obtain more bandwidth than base customers. For an over-dimensioned network, each customer gets his guaranteed rate.

formed experiments from an over-dimensioned (where the sum of the token rates is lower than the Fast Ethernet bottleneck link) to an under-dimensioned network (sum of the token rates is larger than the bottleneck speed). The network is under-dimensioned for base token rates smaller than approximately 10 Mbit/s and over-dimensioned for base token rates above 10 Mbit/s.
We can see that in the 5 to 10 Mbit/s region (corresponding to 45%-100% of network provisioning), there is a clear service differentiation. No explanation has been found for the particularly high throughput of Premium customer 1 for small base token rates. Note that with no service differentiation each customer would get one sixth (approximately 16 Mbit/s) of the bottleneck speed. In an extremely over/under-dimensioned network there is no service differentiation at all, since practically all packets are marked as exceeding (for over-dimensioned network) or as conforming (for under-dimensioned network).
Even if these experiments were run under ideal conditions (the reverse path is not congested, the network is trivial), the very encouraging result is that all customers get their token rate if the network is adequately dimensioned. We can indeed see in Fig. 6 that each customer gets its guaranteed rate for base tokens rates of less than 10 Mbit/s. A customer IP VPN, for example would benefit from such a service.
In our case, 10 Mbit/s corresponds to the transition from an over-dimensioned network to an under-dimensioned network. Our experiments show that network provisioning is key to service differentiation. In our limited test environment, network provisioning is trivial but in larger networks, dealing with a large number of token rates and with a dynamic traffic matrix, network engineering will require sophisticated tools.

Conclusions

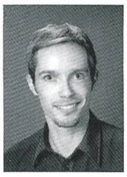
The Internet Engineering Task Force has developed two models for offering network performance guarantees: the Integrated Services model and the Differentiated Services model. The Differentiated Services model is more scaleable than the Integrated Services model and receives the largest interest at this time. Nevertheless, propositions are being made to use RSVP and Integrated Services in the access whereas the backbone uses a Differentiated Services solution. Our results show that routers are now available with mechanisms suitable for offering network performance guarantees. However, routers with suitable mechanisms are not sufficient to offer network performance guarantees in IP networks:
Service providers have to ensure that the guarantees given to the customer are fulfilled. This implies that adequate network management and monitoring tools have to be developed and implemented for provisioning, trouble-shooting and monitoring the offered network performance. When network management and monitoring tools operate successfully with network performance guarantees enabling routers, IP networks will be able to offer loss, delay and throughput guarantees. The transition from IP networks with no performance guarantees to IP network offering network performance guarantees will occur within two years, at least in networks managed by one operator with routers from a single supplier. 9.4

Zusammenfassung

Quality of Service in IP-Netzwerken – bald Realität?

Das aktuelle Internet bietet keine Verlust-, Verzögerungs- oder Durchsatzgarantien an. Solche Netzgarantien, auch verstanden unter «Quality of Service», sind unabdingbar wenn Echtzeitapplikationen wie IP Telefonie und Videokonferenzen übertragen werden. Dieser Mangel an Netzdienstgüte hat das Internet Engineering Task Force (IETF) zum Entwickeln neuer Standards veranlasst, während Hersteller neue Router bauen. Das IETF hat zwei Modelle für Netzdienstgüte entwickelt: das Integrated Services Modell und das Differentiated Services Modell. Das Differentiated Services Modell ist skalierbarer als das Integrated Services Modell und geniesst im Moment das grösste Interesse. Anhand eines Tests zeigen wir ein Beispiel für differenzierte Netzdienstgüte. Aufgrund dieser Entwicklungen gehen wir davon aus, dass die Übertragung über unabhängige IP Netze (1 Netzbetreiber mit 1 Lieferanten) innerhalb der nächsten 2 Jahre mit Verlust-, Verzögerungs- und Durchsatzgarantien erfolgen kann.

Dr. Dominique Moix *received his diploma in physics in 1990 and his Ph. D in applied physics in 1995, both from the ETH Zurich. In 1996, he joined Swisscom Corporate Technology. He has worked in traffic and performance aspects of ATM and IP. He is currently interested in Quality of Service deployment in IP networks.*



Fabien Berger *graduated in Communication Systems from EPFL and Institut Eurecom, France in 1996. He then joined Swisscom Corporate Technology to evaluate pre-MPLS technologies. Since early 1998, he has been working on QoS technologies for IP networks.*

Dr. Fritz Braun *received his diploma in electrical engineering in 1970 and his Ph. D in 1975, both from ETH Zurich. His main activities have been in the areas of digital signal processing, transmission systems, LAN, PABX, traffic engineering and network performance. He is presently working as a project leader in Swisscom Corporate Technology.*