PCP: A Bandwidth Guaranteed Transport Service for IP networks.

Flaminio Borgonovo, Antonio Capone, Luigi Fratta,
Mario Marchese, Chiara Petrioli

Politecnico di Milano

E-mail: {borgonov, capone, fratta, mmarches, petrioli}@elet.polimi.it

ABSTRACT: The increasing demand for a variety of new Internet services with different and possibly stringent QoS requirements (i.e. Internet Telephony, videoconferencing etc.) requires the design of mechanisms to support QoS guarantees. The current solutions proposed in IETF, RSVP (Resource reSerVation Protocol) and Differentiated Services, though suitable for many applications, may result inefficient to support real-time services on a call basis. RSVP is not really scalable and requires substantial changes in the Internet architecture, while Differentiated Services provide guarantees mostly on a static and permanent basis. In this paper we assume an Internet architecture supporting multiple priorities as needed in Differentiated Services, and introduce PCP (Phantom Circuit Protocol), a mechanism that provides a guaranteed bandwidth transport service for circuit oriented connections. PCP includes a fully scalable Call Admission Control (CAC) and operates on a per call basis. Simulation of the protocol performance for CBR (Constant Bit Rate) traffic under various network conditions show the adherence of the mechanism to theoretical expectations.

I. INTRODUCTION

Many efforts have been recently devoted by the Internet community to investigate a transport mechanism capable of guaranteeing demanding Quality of Service (QoS) requirements. The goal is to offer an alternative to carry voice, video and multimedia with respect to classic Telephone/ISDN and ATM networks. The basic problem is how to guarantee bandwidth, delay and packet dropping probability, as requested by voice and video, in a datagram network architecture where the only service is the Best Effort packet transmission.

In this paper we refer to QoS control solutions which are completely implemented at the IP level and do not assume any lower levels QoS guarantee capability. In this context, two different approaches have been proposed in the Internet world: RSVP (Resource reSerVation Protocol)[1, 2] and DS (Differentiated Service) [3, 4, 5, 6].

RSVP is a signaling mechanism among routers and hosts that includes support to "flows" of packets with different QoS and the ability to dedicate end-to-end capacity by means of hop-by-hop resource reservation protocols. In practice, this solution changes the entire network architecture by relying on the virtual circuit connection mechanism, the paradigm of the telephony world, today extended to the B-ISDN. The reservation and signaling procedures are complex and hardly scalable. All network routers in the path of a new connection are involved in the reservation and admission control procedures and must have knowledge of individual incoming flows. Though a more scalable variant has been recently proposed [7], where aggregate traffic is considered and the call acceptance procedures are based on real-time estimates of the available bandwidth at routers, a substantial change in the routers is still required.

The IETF (Internet Engineering Task Force) has therefore proposed an alternative and simpler solution based on Differentiated Services. The basic idea is to use either the IPv4 header TOS (Type Of Service) bits or the IPv6 Traffic Class octet, the "DS byte" to designate the per-hop behaviors (PHB) that packets are to receive. For example, at each router, frames of class A can be forwarded before frames of class B, or frames of class C can be dropped after frames of class D. By carefully aggregating a multitude of QoS-enabled flows into a reasonable number of differentiated services offered by the network it is no longer necessary to recognize and store information

0-7803-5284-X/99/$10.00 © 1999 IEEE.

about each individual flow in the core routers. The network operation still remains purely datagram and scales well. However, the translation of differentiated service at routers into concrete end-to-end guarantees requires call admission control procedures limiting the number of admitted high priority connections. Static solutions have been proposed in [4] but the problem of designing simple and scalable mechanisms enforcing end-to-end bandwidth and delay guarantees on a call basis remains an open problem.

In this paper we introduce PCP (Phantom Circuit Protocol), a mechanism that provides a guaranteed bandwidth service in IP networks supporting service priorities. PCP includes an effective Call Admission Control (CAC) protocol and operates on a call basis, as in circuit switched networks. The CAC procedure only involves network access points and does not require network routers to exchange any call set-up signaling. The protocol therefore conjugates the simplicity and scalability of the DS approach with the RSVP capability of providing end-to-end guarantees on a call basis. Simulation results provided in section IV establish the validity of the proposed approach for CBR (Constant Bit Rate) traffic. However, PCP can be easily extended to support VBR (Variable Bit Rate) traffic as well. PCP behavior under VBR ON/OFF traffic is currently under investigation.

The remainder of the paper is organized as follows. In section II and III the PCP protocols is thoroughly described and discussed. In particular, the basic PCP mechanisms are highlighted in section II, while a possible implementation of the bandwidth test is described and discussed in section III. Simulation results evaluating the protocol performance for CBR traffic under various network scenarios are then presented in section IV. Conclusions are given in section V.

II. THE PCP PROTOCOL

PCP is a mechanism able to provide a bandwidth-guaranteed connection-oriented transport mechanism between network edges using the packet transfer capability provided by a datagram network protocol such as IP. It only requires that each packet in the network belongs to one of the three following priority classes:

- Class 0: (lowest priority) if the packet requests best effort service;
- Class 1: (intermediate priority) if the packet is a probing packet, used in the set-up procedure as defined below;
- Class 2: (highest priority) if the packet belongs to a flow that has been accepted for guaranteed service.

The priority information is carried in the IPv4 TOS or IPv6 DS field and is used by the routers to serve all packets according to a non-preemptive head-of-the-line priority scheme.

To initiate a call, the calling user signals the connection parameters, such as the requested bandwidth and the traffic profile, to its network access point, NODE_A. NODE_A then starts immediately performing an end-to-end CAC in collaboration with the recipient's network access point, NODE_B. The call admission procedure has the purpose to look for the requested bandwidth and to seize it if available. In this phase, the set-up procedure must not affect the QoS of accepted calls.

The key principles on which PCP is based are stated in the following.

The bandwidth availability is assessed by an end-to-end measure performed by network access points NODE_A and NODE_B with no involvement of the network routers. The measure is performed by having NODE_A transmitting probing traffic packets, addressed to NODE_B, according to the traffic profile requested by the user. The recipient NODE_B performs measures on the incoming probing packets stream to verify whether requirements are met and the result is signaled back to NODE_A. If a positive response is received, NODE_A replaces probing packets with user data packets and the transport session between the two access points is used to transfer them to the final recipient.

Note that the transmission of probing packets belonging to class 1 does not steal bandwidth from already established connections (class 2 packets). The probing packets will reach the destination with the desired QoS only if enough network resources (bandwidth), not used by class 2 traffic, are available. The CAC proper operation requires that the replacement of probing packets with data packets occur with no interruption and that the users with accepted calls continuously use the bandwidth until connection release. The bandwidth allocated to a connection is automatically released when data packets are no longer transmitted by NODE_A.

Standard feedback and time-out procedures are adopted to coordinate operation between NODE_A and NODE_B.

PCP operation is bandwidth-measure, rather than network state, based. This characteristic makes the mech-

anism scalable, but requires that, in order to provide a reliable guaranteed bandwidth service, the measure at the call set-up truly reflects the current available bandwidth. To enforce this property it is important that the measure period be long enough to capture the traffic dynamics. In this paper we only consider CBR traffic, where all sources continuously transmit at the peak rate. In this case it is requested that the set-up period be tuned to capture a sufficient number of packets of the slowest connection in the network, as explained in the next section. The applicability of PCP with VBR traffic, whose effectiveness is currently under investigation, requires the definition and the measure of a suitable average bandwidth instead of the peak one.

III. BANDWIDTH MEASURE

In this section we describe a possible measure on the sequence of the probing packet interarrival times X_i to decide whether a call should be accepted.

In the case of CBR traffic the bandwidth to be verified is represented by the constant rate B at which packets are transmitted. Unfortunately, statistical multiplexing introduces some jitter in the packet inter-arrival time X_i at NODE$_B$, with $E[X_i] = \frac{1}{B}$ if the bandwidth is available along the path and $E[X_i] > \frac{1}{B}$ otherwise. Therefore, the problem of assessing the existence of the required bandwidth can be reduced to a statistical test.

The average $E[X_i]$ can be estimated by the sample average

$$M_X = \frac{\sum_{i=1}^{N} X_i}{N} \tag{1}$$

on $N+1$ received probing packets. However, hypothesis $E[X_i] > 1/B$ can not be effectively tested against $E[X_i] = 1/B$, due to the variance in M_X. In fact, if we set the discriminating threshold $D = \frac{1}{B}$, too many calls that could be accepted are rejected. If we set the discriminating threshold $D > \frac{1}{B}$, the reverse situation can occur in which a call is accepted even if the required bandwidth exceeds network allowance.

To perform a more efficient test, we change the set up procedure and assume that probing packets are transmitted by NODE$_A$ at a rate $r = B(1+\delta)$, where δ represents a safety factor. This modification does not affect the correct operation of the protocol, but allows to introduce a statistical test of significance between hypothesis I $E[X_i] = 1/r$ (the network has plenty of bandwidth) and hypothesis II $E[X_i] \geq 1/B$ (the most unfavorable case is when the available bandwidth is just below B). The call is accepted whenever the sample interarrival mean $E[M_X]$ is less than a threshold $D_\beta \leq \frac{1}{B}$, which is dynamically computed and set to make negligible the type II error probability β. i.e., the probability that the call is accepted when $E[X_i] \geq 1/B$ is true. We have evaluated the sample variance as

$$S_X^2 = \frac{\sum_{i=1}^{N}(X_i - M_x)^2}{N}. \tag{2}$$

and set the threshold to

$$D_\beta = 1/B - 3\sqrt{\frac{S_X^2}{N-1}} \tag{3}$$

Since the values X_i and X_{i+1} are negatively correlated, expression (2) leads to an estimate of the variance that has an average greater than the true value. This further reduces the probability of a type II error.

Since the measure must verify the existence of bandwidth in a time multiplexed environment, it must be long enough to capture the bandwidth variations of the slowest traffic. For example, if we assume that the lowest bandwidth corresponds to 32 Kb/s with 640 bit packets, the packet inter-transmission time is 20 ms. Thus, a measure period of $1-2$ seconds is needed to collect a reasonable number $(50-100)$ of samples.

The value of δ affects the type I error probability, i.e. the probability of rejecting a call that could be accepted. This error decreases as δ increases, so that the fraction of unnecessary rejected calls stabilizes. In the cases we have simulated, and reported in the next section, we have set $\delta = 0.2$. However, we have found that the procedure we have described is not critical with respect to a wide range of parameters' values.

IV. SIMULATION RESULTS

In order to evaluate the effectiveness of the proposed CAC, we studied the protocol behavior with CBR traffic. In particular, both 32Kb/s and 1Mb/s CBR connections were considered, modeling respectively IP voice and video traffic. We also assumed fixed-size packets of 1000 bits.

The presentation of the simulation results is divided into two parts. First, we show the accepted vs. offered load and the packet delay referring to a single link scenario. Then, results are presented showing the protocol behavior in a multi-link, homogeneous network.

In the simulations for the single link, we have considered a dynamic scenario where calls, generated according

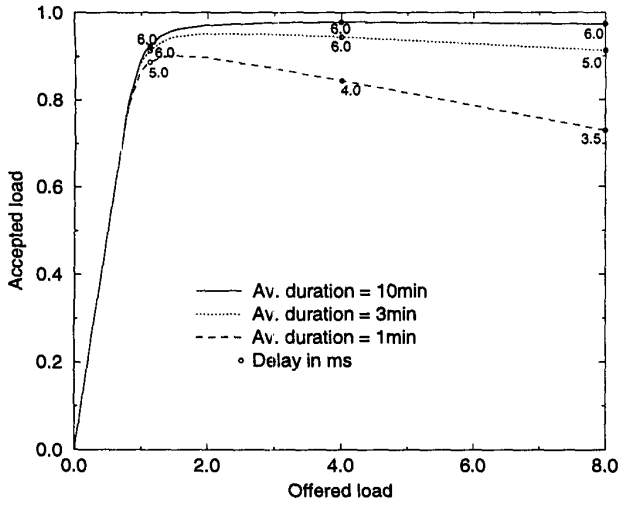

Figure 1: *Accepted vs. Offered traffic. 50000 32Kb/s CBR connections with mean duration* $M_c = 1min, 3min, 10min$ *are dynamically generated according to a Poisson process and transmitted over a 2Mb/s link.* $T_{acc} = 2s$ *and* $\delta = 0.2$.

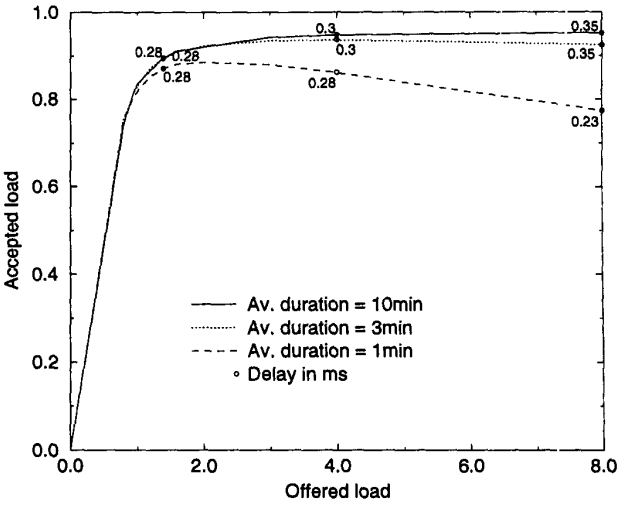

Figure 2: *Accepted vs. Offered traffic. 50000 1Mb/s CBR connections with mean duration* $M_c = 1min, 3min, 10min$ *are dynamically generated according to a Poisson process and transmitted over a 25Mb/s link.* $T_{acc} = 2s$ *and* $\delta = 0.2$.

to a Poisson process, have an exponentially distributed duration, with mean M_c =1 min, 3 min, 10 min. Based on the qualitative criteria previously introduced, T_{acc} has been chosen equal to $2s$, corresponding to 64 packet inter-arrivals of the slowest connections, while δ has been set to 0.2.

The accepted vs. offered traffic is plotted in Figures 1 to 3. In particular, Fig. 1 and 2 analyze homogeneous traffic scenarios with 32Kb/s and 1Mb/s calls respectively. The effect of multiplexing voice and video traffic on the same channel is then displayed in Fig. 3 , where curves showing the accepted traffic for the two types of traffic are also plotted. The 99−percentile of the delay distribution expressed in msec is also shown in the figures for selected samples. To assure confident results, 50000 connections were simulated for each sample.

All curves show that as the offered load increases, the accepted load grows up to a value, which represents the channel utilization achievable by the proposed technique. From this point on the accepted traffic monotonically decreases with the offered traffic because of the contention caused by the increased number of overlapping call set-ups. For a given offered load and T_{acc}, the number of overlapping set-ups decreases with M_c, as it is shown by the increased channel utilization of the 10 min case over the 3 min and 1 min cases. The increased 32Kb/s accepted traffic over the 1Mb/s accepted traffic displayed in Fig. 3 reflects the low probability of finding 1.2Mb/s available bandwidth at high load.

In all the considered scenarios the probability p_f of a call being accepted when not enough bandwidth is available was measured equal to zero.

Finally we observe that the 99−percentile of the delay is very low, even if, as expected, it increases as the accepted traffic increases.

To validate the measuring mechanism and to measure the delays in the most critical case the protocol behavior is therefore also simulated in a statically loaded network. To this purpose, a 9 node network, interconnected by a unidirectional and homogeneous ring structure is considered. Connections with fixed rate 1Mb/s (32Kb/s) are generated at every node, while links have the same capacity equal to 50 or 100 times the connection rate. The system is loaded by connections sequentially generated so that only one call at the time is in set-up. At first, 60% of the link capacity is loaded by 1-hop connections, then, 9-hop connections, with randomly generated sources, are offered up to 5 times the residual capacity. 1 hop connections are introduced only to limit the pipeline effect and their performance are not evaluated. Measures have been obtained over 100 independent loadings of the network. The number of accepted 9−hop connections, and the 99−percentile of the packet delays are summarized in TABLE 1.

TABLE 1. THROUGHPUT AND DELAY PERFORMANCE OF PCP 9-HOP CONNECTIONS.

connection rate	link capacity	
1 Mb/s (32 kb/s)	50Mb/s (1.6 Mb/s)	100Mb/s (3.2 Mb/s)
1-hop conn. number	30	60
9-hop conn. number	19	39
99-percentile of delay (msec)	1.12 (35)	0.77 (24)

We observe that the proposed protocol achieves high links utilization while not exceeding the links capacity. The 99−percentiles of the delay are extremely low and in the considered scenarios never exceed 1.12 msec for video traffic and 35 msec for voice traffic.

Finally note that, even if PCP prevents guaranteed traffic from fully utilizing the network bandwidth, the full resources utilization is still achievable when best effort (class 0) traffic is also considered.

V. CONCLUSION

In this paper we have introduced the PCP protocol, which is able to provide a guaranteed bandwidth transport service in IP networks supporting multiple service priorities. PCP includes a simple, distributed, fully scalable Call Admission Control (CAC) which is based on end-to-end measurements and operates on a per call basis. To assure that the set-up procedure does not affect established calls QoS, a priority based forwarding mechanism is adopted and set-up packets are transmitted at a lower priority than data packets.

The performances of PCP, evaluated through simulation for CBR traffic and under various network conditions, have been reported and show the complete adherence of the mechanism to the theoretical expectations.

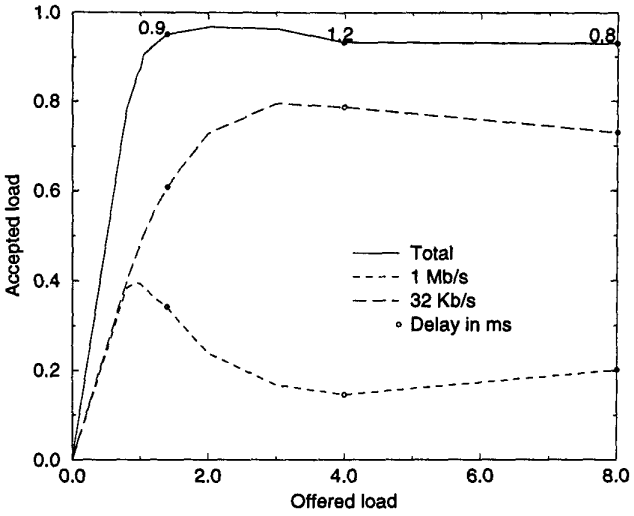

Figure 3: *Accepted vs. Offered traffic over a single 25Mb/s link in a multiple rate scenario. 50000 CBR connections with rate either 1Mb/s or 32Kb/s and mean duration $M_c = 3min$ are dynamically generated according to two Poisson processes so that the global offered traffic is fairly shared among the two types of traffic. $T_{acc} = 2s$ and $\delta = 0.2$.*

REFERENCES

[1] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, "Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification", *IETF Request For Comments 2205*, Sep. 1997.

[2] P.P White, "RSVP and Integrated Services in the Internet: A Tutorial", *IEEE Communication Magazine*, vol. 35, no. 5, May 1997, pag. 100.

[3] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An architecture for Differentiated Services", *IETF Internet Draft*, Oct. 1998.

[4] Y. Bernet, J. Binder, S. Blake, M. Carlson, S. Keshav, E. Davies, B. Ohlman, D. Verma, Z. Wang, W. Weiss, "A framework for Differentiated Services", *IETF Internet Draft*, Oct. 1998.

[5] J. Heinanen, "Assured forwarding PHB", *IETF Internet Draft*, August 1998.

[6] V. Jacobson, "Expedited forwarding PHB", *IETF Internet Draft*, August 1998.

[7] W. Almesberger, T. Ferrari, J. Le Boudec, "Scalable resource reservation for the Internet", *IETF Internet Draft*, Nov. 1997.