Noname manuscript No.
(will be inserted by the editor)

RCFT: A Termination Method for Simple PCN-Based Flow Control

Frank Lehrieder · Michael Menth

Received: date / Accepted: date

Abstract Pre-congestion notification (PCN) conveys information about load conditions in Differentiated Services IP networks to boundary nodes. This information is currently used for admission control and flow termination. Flow termination complements admission control, e.g., in case of failures when admitted traffic is rerouted and causes overload on backup paths. Existing approaches for PCN-based admission control and flow termination operate on ingress-egress aggregates and rely on a signalling protocol that regularly reports measured PCN feedback from all egress nodes to all ingress nodes. However, this signalling protocol is neither defined nor available, and the methods have also other intrinsic shortcomings that result from their operations on ingress-egress aggregates.

While there is already a PCN-based admission control method that works without additional signalling of measured PCN feedback, a solid flow termination method with that property is still missing. In this paper we present the novel regular-check-based flow termination method (RCFT). It does not rely on measured PCN feedback, fills the identified gap, and allows for a PCN architecture without signalling of measured feedback. We explain RCFT in detail and investigate its termination behavior under various conditions. Moreover, we study the use of PCN-based flow control for on/off traffic. These results are of general nature and apply to any system using PCN-based flow termination.

Frank Lehrieder
University of Würzburg, Germany
E-mail: lehrieder@informatik.uni-wuerzburg.de

Michael Menth
University of Tübingen, Germany
E-mail: menth@informatik.uni-tuebingen.de

Keywords quality of service – pre-congestion notification – admission control – flow termination

1 Introduction

Due to the increasing fraction of real-time traffic (voice and video) in today's networks, Internet service providers and manufacturers have recognized the need for admission control (AC) in some parts of the Internet. AC limits the rate of high-priority traffic by explicitly admitting or blocking new flows to avoid congestion. A standardized method is flow reservation using the Resource reSerVation Protocol (RSVP) [1]. It is considered complex because it requires per-flow states in each router along a path whose quality of service (QoS) needs to be protected.

In response to this shortcoming, the Internet Engineering Task Force (IETF) recently proposed pre-congestion notification (PCN) to support AC and flow termination (FT) in Differentiated Services (DS) IP networks. In spite of AC, overload may occur in unexpected situations, e.g., when links or nodes fail and traffic is rerouted over alternative path on which congestion may occur. Then, FT terminates some admitted flows to restore a controlled load condition [2].

PCN is applied on a per-domain basis. Routers meter the load of PCN traffic on each link of a domain and re-mark packets if link-specific rate thresholds are exceeded. In the currently standardized approach [3], the egress nodes continuously measure the rates of differently marked traffic per ingress-egress aggregate (IEA) and regularly report them to the corresponding ingress nodes which use this PCN feedback to perform AC and FT.

This method has several shortcomings. The egress node needs to classify every PCN packet for the measurement process. The ingress node needs to find the appropriate IEA for a new flow request to perform AC and the proposal is unable to deal with multipath routing [4, 5]. Last but not least, there is significant signalling overhead between all ingress and egress nodes of a PCN domain and the signalling protocol is only under discussion [6] but not yet available.

In this paper, we present a novel PCN-based FT algorithm that checks each admitted PCN flow in regular intervals for the marking of its last packet for potential termination. We call the method regular-check-based flow termination (RCFT). A salient feature of RCFT is the fact that it does not require measured PCN feedback. Together with marked-signalling-based AC (MSAC) [4, 7] it allows for PCN-based flow control that is clearly simpler than the current experimental specification [3] and works well in networks with multipath routing. We analyze the termination behavior of RCFT by means of simulations, show that it works well under various conditions, and give recommendations for its configuration. Moreover, we study the use of PCN-based flow control in the presence of on/off traffic; the results of these experiments are also applicable to PCN-based systems with other FT methods.

Section 2 gives an introduction to PCN. Section 3 reviews related work about PCN and flow termination in general. Section 4 presents the new RCFT and shows how it helps to build a system for PCN-based AC and FT without signalling measured PCN feedback. Section 5 investigates the termination behavior of RCFT under various conditions and evaluates the use of PCN-based flow control in the presence of on/off traffic. Section 6 concludes this paper. The appendix contains a glossary to facilitate reading of the paper.

2 Pre-Congestion Notification

We illustrate the general concept of PCN [8] and its application for AC and FT purposes. We summarize the current architecture [3] and point out its shortcomings.

2.1 The Concept of PCN-Based AC and FT

PCN protects the QoS of high-priority flows in a single DS domain; these flows are called PCN flows. Their packets are marked as PCN [9] and are preferentially forwarded compared to other traffic. PCN introduces an admissible and a supportable rate threshold ($AR(l)$, $SR(l)$) for each link l of the network. These thresholds imply three different pre-congestion states on a link as illustrated in Figure 1. If the PCN traffic rate $r(l)$ is below $AR(l)$, there is no pre-congestion and further flows may be admitted. If the PCN traffic rate $r(l)$ is above $AR(l)$, the link is AR-pre-congested and no further flows should be admitted. If the PCN traffic rate $r(l)$ is above $SR(l)$, the link is AR- and SR-pre-congested and some already admitted flows should be terminated.

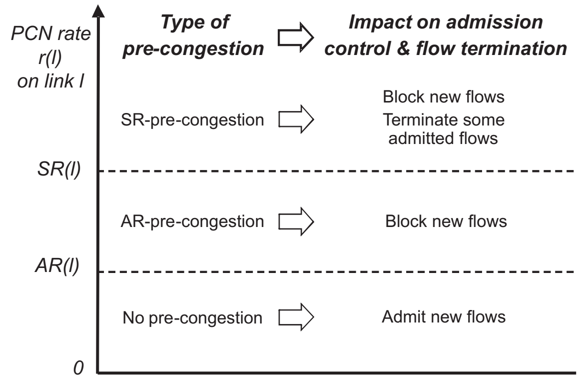


Fig. 1 The admissible and the supportable rate $(AR(l), SR(l))$ define three pre-congestion states with respect to the PCN traffic rate $r(l)$ on a link.

PCN-based AC assumes that PCN flows request admission before sending traffic through the PCN domain. This request is either admitted or blocked by a decision point of the domain. The admission requests are signalled by another protocol which may be, e.g., RSVP. In that case, only the decision point processes RSVP messages by proxy for the entire DS domain so that the remaining DS domain can be unaware of individual PCN flows. However, appropriate policers at the ingress nodes still need to be configured by the decision point so that packets of admitted PCN flows can enter the DS domain. With path-coupled resource signalling, the admission request travels the same path as future data packets. In such a scenario, the decision point is collocated with the ingress or the egress node. With path-decoupled resource signalling, the decision point may be implemented as a centralized node and contacted by all admission requests for the domain. An example is the Resource and Admission Control Function (RACF) in Next Generation Networks [10]. These issues are discussed in more detail in [11]. In the remainder, we focus only on path-coupled signalling.

2.2 The CL Method

We explain the operation of the "Controlled Load" (CL) method for PCN-based AC and FT [3].

2.2.1 Metering and Marking

PCN packets enter a PCN domain with a “not-marked” (NM) codepoint and are appropriately re-marked by PCN nodes in case of pre-congestion. Two different marking algorithms are in use [12]: threshold marking and excess traffic marking. The reference rate of threshold marking is set to $AR(l)$. If the PCN traffic rate $r(l)$ exceeds that value, the threshold marker re-marks all NM-packets to “threshold-marked” (ThM). The reference rate for excess traffic marking is set to $SR(l)$. The excess traffic marker leaves NM- and ThM-traffic of that rate untouched and re-marks all other traffic to “excess-traffic-marked” (ETM). Both marking algorithms work with a certain tolerance that is specified by token bucket parameters. Since the load thresholds are usually set to a lower value than the link bandwidth, re-marked PCN packets indicate an increased load condition within the PCN domain long before congestion occurs so that this technique is called pre-congestion notification. The encoding of NM, ThM, and ETM in the header of IP packets is described in [9].

2.2.2 Distribution of PCN Feedback

The decision points in the CL method are collocated with the ingress node. Many operations are performed per IEA which is the ensemble of all flows sharing a common ingress and egress node. The egress node measures the rate of NM-, ThM-, and ETM-traffic (NMR, TMR, and EMR) per IEA in regular intervals of about 200 ms, and signals them as PCN feedback to the ingress nodes.

2.2.3 Admission Control

When the decision point receives a PCN report for an IEA, it calculates the fraction of re-marked packets, the congestion level estimate $CLE = \frac{TMR+EMR}{NMR+TMR+EMR}$. If that value is larger than a configured CLE limit L_{CLE}, the PCN admission state for that IEA is set to “block”; otherwise, the PCN admission state is set to “admit”. The latter is also done if the IEA is empty. Depending on the PCN admission state, the decision point admits or blocks new admission requests.

2.2.4 Flow Termination

When the decision point detects that the rate of ETM-traffic is larger than zero ($EMR > 0$), some flows of the corresponding IEA should be terminated. To that end, the decision point requests the rate of admitted PCN traffic from the ingress node (ingress rate, IR). The ingress node measures that rate and returns it to the decision point. Then, the decision point calculates the rate of traffic to be terminated as $TR = IR - NMR - TMR$ and chooses an appropriate set of flows with that overall rate for termination. This is not a trivial task as traffic descriptors used for policing usually overestimate the flow rate so that several termination steps may be required. Sufficient time should elapse between successive termination steps to avoid that more traffic is terminated than needed (overtermination) [5]. More intuitive termination methods are possible, but the presented method assures that sufficient traffic is quickly terminated even in case of heavy traffic loss.

2.3 Shortcomings of the CL Method

The CL method suffers from several shortcomings. It does not properly work with multipath routing as the signalled PCN reports from egress to ingress nodes do not differentiate the feedback from individual paths. As a result, underadmission may occur, i.e., too little traffic may be admitted [4]. In a similar way, over- and undertermination may occur [5]. While mapping PCN packets to individual flows seems to be a standard operation, mapping flows to specific IEAs is difficult in general environments. A special challenge for the ingress node is to map an admission request for a new flow to its corresponding IEA as the corresponding ingress node within the DS domain is generally not known. The CL method requires extra signalling to convey PCN reports from all egress nodes to all ingress nodes, about 5 times per second. This burdens ingress and egress nodes as well as the network with additional signalling load. A protocol option can reduce this signalling load, but only in the absence of pre-congestion [3]. Moreover, the signalling protocol for the transport of measured PCN feedback is still under discussion [6] and not yet available.

3 Related Work

We give an overview of related work in the PCN area and point out activities in flow termination.

3.1 Activities in Pre-Congestion Notification

Typical measurement-based admission control (MBAC) measures the rate of admitted traffic and takes admission decisions on that basis. With PCN, feedback is provided by markers inside the network; in particular, the markers are configurable so that early marking is provided before congestion occurs. Early work in this area can be found in [13] and [14].

An overview of PCN including a multitude of AC and FT mechanisms is given in [15]. In [16], a high level summary about a large set of simulation results for PCN-based AC and FT was provided and it was shown that these methods work well in most studied cases.

The authors of [17] propose an autonomic PCN-based AC algorithm optimized for video services in multimedia access networks and evaluate it with typical video traffic. As the shim header in Multiprotocol Label Switching (MPLS) provides even fewer codepoints than the IP header, the authors of [18] proposed an encoding scheme for threshold marking and excess traffic marking in a single codepoint using the frequency of re-marked packets to interpret the case-specific meaning of the codepoint. They compared the perceived quality of experience of voice flows with probe-based AC which is very similar to marked-signalling based AC. Probe-based AC was also implemented in a different context using the Session Initiation Protocol (SIP) [19]. The problem of the shortage of codepoints has been addressed in [20] by an alternative encoding scheme called packet-specific dual marking (PSDM). PSDM allows tunneling with legacy equipment while the PCN encoding [9] in the standardization process requires special tunneling rules in a PCN domain [21]. The PCN variant suggested in this paper works with both encoding options.

In [5], multiple options for PCN-based flow termination using IEA-based measured feedback were investigated. Overtermination due to different round trip times was studied in [22]. In [23], the concept of PCN-based marked flow termination (MFT) was proposed and evaluated. MFT does not need to signal PCN feedback from egress nodes to ingress nodes, but the presented mechanisms were rather complex and less robust against different flow rates than regular check termination that we introduce and study in this work. In [24], MFT was adapted to the Single Marking architecture in [25] and its performance was evaluated. Overtermination due to multiple bottlenecks was investigated in [26] for various termination methods. The study in [27] gives recommendations for the setting of admissible and supportable rate thresholds for resilient networks so that admitted flows are not terminated after rerouting in case of single link failures. Furthermore, it studied how link weights should be set in IP networks to maximize admissible traffic rates.

3.2 Flow Termination in Other Contexts

A similar mechanism to flow termination was included in Multi-Level Precedence and Preemption (MLPP), which was specified in year 1993 by the ITU [28]. MLPP consists of two parts: precedence and preemption. Precedence associates each call that is made in the network with a priority level. With preemption, higher-priority calls can make lower-priority calls to be torn down in case of scarce network resources. The specification states that users of terminated calls should receive a notification when their call is terminated. A number of standardization bodies include this feature today in the network specifications. Besides the specification for the Integrated Services Digital Network (ISDN) [29] it was also defined more recently for the Session Initiation Protocol (SIP) in [30] by the Internet Engineering Task Force (IETF). In addition, also wireless networks such as GSM, UMTS and LTE can use this feature networks [31–33], where it is called enhanced Multi-Level Precedence and Preemption (eMLPP).

MLPP is mostly studied in the area of military communications. For example, [34] proposes algorithms for military IP networks to implement MLPP in packet based switching networks. In addition, [35] investigates how measurement-based admission control can be used for MLPP in wireless ad-hoc networks. Fineberg [36] analyzes how MLPP can be implemented using the Differentiated-Services-aware traffic engineering (DS-TE) in MPLS networks. Also Shan and Yang [37] improve resource sharing by means of preemption in a Differentiated-Services-aware environment.

Preemption of flows or calls in MPLS networks is further investigated in a number of studies. Some works [38, 39] propose preemption policies for MPLS networks and evaluate their performance. Other studies investigate bandwidth constraint models [40] or bandwidth allocation strategies with preemption [41]. In addition, some more theory-oriented studies exist on the optimal set of calls to be terminated and the computational complexity of algorithms to determine these flows [42, 43]. Concretely, Dogar et al. [43] show that it is an NP-complete problem to minimize the number of preempted flows and the preempted bandwidth in multi-class networks.

In contrast to PCN, these mechanisms for flow preemption are specifically targeted to guarantee that high priority calls can immediately be established even in case that the available resources are not sufficient. The flow termination function of PCN targets other scenarios. It is intended to tear down some of the already admitted flows in a controlled way if unexpected events happen in the network such as router or link outages.

The flow termination mechanisms of PCN can respect flow priorities but it is not their main objective. In contrast, a controlled load situation [3] should be quickly restored.

4 A New Flow Termination Method for Simple PCN-Based Flow Control

In this section we present RCFT as a new method. We explain its benefits and compare it with other FT methods to underline its novelty. Then, we review a PCN-based AC method that excels with similar advantages. Eventually, we propose the combination of this AC method and RCFT as a simple architecture for PCN-based flow control.

4.1 Regular-Check Based Flow Termination (RCFT)

We describe the operation of RCFT, point out its benefits, and compare it with other FT methods.

4.1.1 Operation of RCFT

With RCFT, the egress nodes of a PCN domain check all admitted flows in regular intervals (check intervals) of duration Δ_{Check}. If the marking of a flow's most recent packet was excess-traffic-marked, the egress node terminates that flow. This may be done by sending an appropriate end-to-end signalling message, e.g., PATHTEAR and RESVTEAR when RSVP is used. We call this termination method "regular-check based flow termination" (RCFT). The egress node initializes this process by setting the timer to the parameter Δ_{Check} when it admits the new flow. The checkpoints for all flows of an egress node may by synchronized or intentionally desynchronized to smooth the control overhead for the egress node over time.

4.1.2 Benefits of RCFT

RCFT excels by the fact that it uses excess traffic marking (see Section 2.2.1) which has already been standardized. It does not need measurement of PCN feedback so that the egress node does not need to map PCN packets to IEAs which is not a trivial operation. RCFT does not require a signalling protocol which makes it simpler to deploy than the recently standardized CL method [3]. RCFT works well in networks with multipath routing since only such flows are terminated for which packets are carried over a SR-pre-congested link. Last but not least, RCFT does not need the knowledge of flow rates for proper operation.

4.1.3 Comparison of RCFT with Other Flow Termination Methods

RCFT can be classified as a marked flow termination (MFT) method as it terminates flows based on marked packets only, without the measurement of PCN feedback. In [23] we proposed three different MFT methods: MFT-MFR, MFT-IF, and MFT-IEA that we review in the following.

MFT-MFR requires excess traffic marking with marking frequency reduction (MFR) which is not a standardized method. With MFT-MFR the egress node terminates a PCN flow as soon as one of its packets is marked as ETM. PCN flows with larger packet rates are terminated with higher probability than flows with smaller packet rates, which is clearly unfair.

MFT-IF uses plain excess traffic marking. Upon admission of a flow, the egress node initializes a credit counter for that flow. This counter is decremented by the size of each packet that has been received by the egress node for this flow with an ETM-mark. If the counter becomes negative, the flow is terminated. This method suffers from the fact that the egress node needs to know the flow rate for an appropriate initialization of the credit counter in order to avoid unfair termination. Nevertheless, the method suffers from the fact that long-time flows suffer a higher termination probability than short-time flows. MFT-IEA is an extension of MFT-IF to IEAs, i.e., there is a credit counter per IEA and not per flow. This improves some investigated performance metrics, but makes the algorithm more complex, in particular as it requires that PCN packets are mapped to IEAs. Thus, all MFT methods of [23] are either unfair to PCN flows with high packet rates or to long-time flows. This is not the case for RCFT.

The termination methods presented in [5] are classified as measured rate termination (MRT). The FT methods in both the CL and SM PCN architecture [3, 25] fall in that category. Egress nodes maps PCN packets to IEAs, measure the rates of differently marked packets per IEA, and signal this PCN feedback to PCN ingress nodes or other decision points. Thus, RCFT is significantly different from MRT methods and simpler.

Since marked flow termination methods terminate only flows with excess-traffic-marked packets, they cope with multipath routing by design because they terminate only flows whose paths run over SR-pre-congested links. This is different with measured rate termination methods. To cope with multipath routing, they require that information about recently marked flows is signalled to the decision points to terminate only flows that contribute to SR-pre-congestion.

4.2 A Simple Admission Control Method

Probe-based AC (PBAC) with implicit probing has been presented and evaluated in [4]. Then it was proposed in IETF under the name marked-signalling-based AC (MSAC) [7] since probing sometimes implies the generation of extra traffic which is not required for this AC method. We assume that resource reservation and admission requests for flows are performed on the basis of path-coupled signalling protocols such as RSVP. We first present basics about RSVP and then explain MSAC.

4.2.1 RSVP Basics

RSVP is a complex resource reservation protocol with multiple features. We review only basic operations that are relevant in our context.

To set up a reservation, a receiver requesting a media stream or a proxy thereof indicates the sender to set up a reservation. The sender or a proxy thereof then sends a PATH message to the receiver whereby PATH states are installed on all intermediate RSVP-capable hops. They record in particular previous hop information. If the path is unavailable for some reason, a node may return a PATHERR message to inform the sender that the path cannot be set up. When the receiver or a proxy thereof eventually receives a PATH message, it responds with a RESV message requesting the needed resources in all hops along the path and installing a RESV state with flow-related information. To this end, the RESV message is forwarded in the reverse direction of the future media stream using the previous hop information in the PATH states. When a node receives a RESV message for the first time, it performs AC for that flow with respect to downstream resources. If it succeeds, it passes the RESV message to the next upstream RSVP neighbor, otherwise it returns a RESVERR message to indicate that the reservation has failed. When the RESV message reaches the sender, the reservation is established. RSVP implements the soft state concept, i.e., PATH and RESV states automatically expire after some time unless they are refreshed by regular PATH and RESV messages. In addition, any intermediate node on the path may send PATHTEAR and RESVTEAR messages to explicitly and quickly terminate a reservation.

4.2.2 Marked-Signalling Based AC (MSAC)

We explain MSAC using RSVP as an example, but the concept can also be implemented with other path-coupled resource signalling protocols.

MSAC requires that only ingress nodes and egress nodes in a PCN domain act as RSVP-capable nodes for incoming and outgoing traffic. Ingress nodes classify PATH messages as PCN traffic and label them as not-marked PCN traffic. As a result, PATH messages are re-marked if any kind of pre-congestion occurs on the path from ingress to egress within the PCN domain. The egress node forwards this PATH message if it belongs to an already established reservation or if the packet is still not-marked. Otherwise, it returns a PATHERR message to indicate that the reservation across the PCN domain cannot be set up. Thus, the ingress node receives a corresponding RESV message only when the network is not pre-congested. Therefore, it admits new reservations by default when it receives new RESV messages. Thus, the actual admission decision is taken by the egress node by either forwarding first PATH messages or returning PATHERR messages.

MSAC copes well with multipath routing under the condition that RSVP messages and data packets of the same flow are carried on the same path. Then, flows are admitted or blocked depending on the load conditions of their prospective paths. This is different with CL's AC method: a flow may be already blocked if a single path belonging to an IEA is pre-congested although the flow's traffic would be carried over a different, non-pre-congested path [4].

4.3 A Simple Architecture for PCN-Based Flow Control

We suggest to combine MSAC and RCFT to constitute a simple architecture for PCN-based flow control that eliminates some shortcomings of the recently standardized CL architecture [3]. Our proposal uses the same metering and marking algorithms as the CL architecture [3]. The CL architecture measures PCN feedback and needs to map PCN packets to IEAs which can be a difficult task depending on the underlying network architecture. Our proposed solution avoids this difficulty. The CL architecture requires a signalling protocol to convey PCN feedback from egress nodes to decision points. This protocol for PCN signalling is still to be designed, adds overhead in terms of traffic load, and represents another source of failure. Our proposal gets along without such a signalling protocol. It is known that the CL architecture does not perform well in networks with multipath routing. The AC algorithm of the CL architecture may lead to underadmission in case of multipath routing [4], and its FT method requires that egress nodes explicitly signal a set of recently marked flows to decision points for termination purposes. Otherwise, wrong flows may be terminated which may cause significant overtermination [5]. Both MSAC and RCFT do not suffer from these problems.

5 Performance Study

In this section, we first describe our simulation setup and then evaluate the termination behavior of RCFT. In particular, we investigate PCN-based flow control in the presence of on/off traffic which is of interest for the configuration of general PCN-based systems.

5.1 Simulation Setup

We consider a single link and assume sudden SR-pre-congestion as it appears for instance due to rerouted traffic when fast failover mechanisms are used. The inter-arrival time A and the packet size B of the flows in our simulations are deterministic. If not mentioned differently, they have average values of $E[A] = 20$ ms and $E[B] = 200$ bytes such that the traffic rate of corresponding flows is $E[R] = 80$ kbit/s[1]. To avoid simulation artifacts due to marking synchronization of periodic traffic, we add an equally distributed random delay of up to 1 ms to the theoretic arrival instant of every packet. This traffic model is realistic because realtime applications send traffic periodically, but packets arrive at the bottleneck link with some jitter.

The supportable rate of the considered bottleneck link l is $SR(l) = 8$ Mbit/s and the initial SR-overload, i.e., the traffic rate by which the supportable rate $SR(l)$ is exceeded is 100%. Hence, the initial number of flows is $n = 200$, but only $n = 100$ PCN flows can be supported. We simulate the time-dependent PCN rate $r(t)$ on the bottleneck link to study the termination process. We assume a termination delay of $D_T = 200$ ms which is the time that elapses between the termination trigger of the egress node until the egress node does no longer see packets of the terminated flow. It comprises at least one round trip time between egress and ingress node plus some processing delay.

We use a custom-made Java tool to simulate the PCN rate $r(t)$ to illustrate the termination behavior. This rate is calculated based on 50 ms long measurement intervals. We perform multiple experiments and report average results for the termination behavior in our figures. We run so many simulations that the 95% confidence intervals for the PCN rate values $r(t)$ are smaller than 1%. However, we omit them in the figures for the sake of easier readability[2].

[1] $E[X]$ is the mean and $c_{var}[X]$ the coefficient of variation of a random variable X.

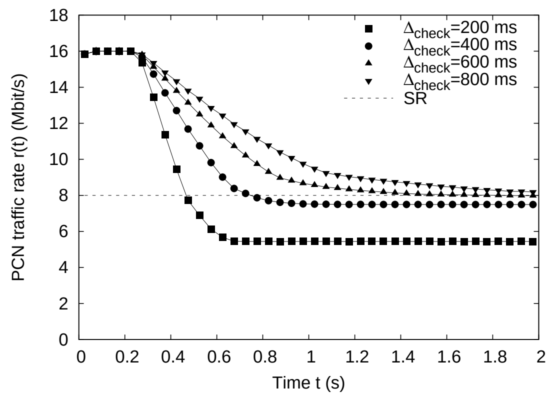


Fig. 2 Impact of the check interval Δ_{Check}.

5.2 Impact of the Check Interval Δ_{Check}

We investigate the impact of the check interval Δ_{Check} by varying this value from D_T to $4 \cdot D_T$ ($D_T = 200$ ms). Figure 2 shows the time-dependent PCN rate $r(t)$ on the bottleneck link when termination starts. A value of $\Delta_{Check} = D_T$ leads to significant overtermination, $\Delta_{Check} = 2 \cdot D_T$ still causes some overtermination, $\Delta_{Check} = 3 \cdot D_T$ leads to fast termination without overtermination, and $\Delta_{Check} = 4 \cdot D_T$ just slows down the termination process. Therefore, we use $\Delta_{Check} = 3 \cdot D_T$ as standard value in the remainder of the paper. Under these conditions, most of the overload is removed after two check intervals.

5.3 Comparison with CL's Termination Method

For $\Delta_{Check} = 3 \cdot D_T$, RCFT is similarly fast as CL's termination method. Depending on the configuration of the meters and markers and the current load, it takes 200 ms or more until PCN nodes start marking traffic in case of pre-congestion [4]. In the CL approach, the egress node requires up to 200 ms to finalize the PCN report when re-marked packets are visible at the egress node. The report is signalled to the decision point which requests from the ingress node the ingress rate IR whose measurement requires another 200 ms. Before traffic can be terminated, it takes up to two measurement intervals plus signalling time from the egress

[2] Even in case of strictly periodic traffic, i.e., the inter-arrival times and the sizes of the packets are constant, different runs produce different results because the first transmission of each flow within the first inter-arrival time after simulation start is random.

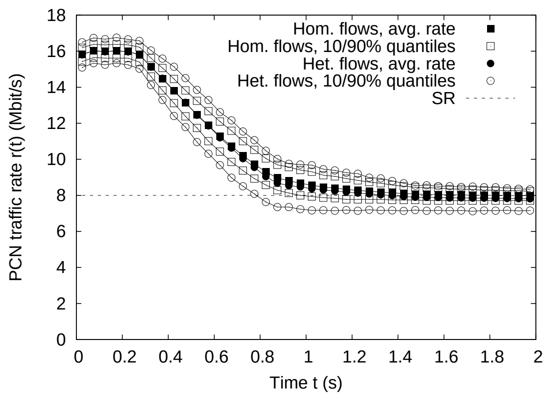


Fig. 3 Impact of flow rate variability.

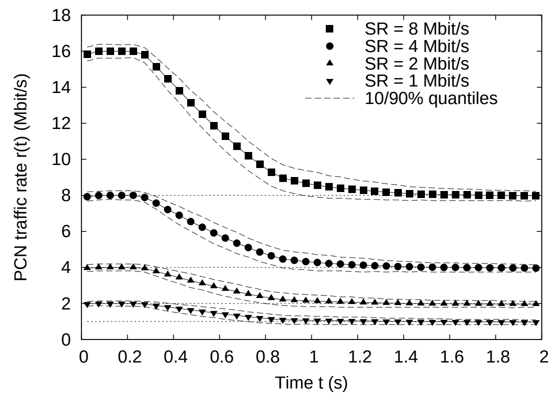


Fig. 4 Impact of traffic aggregation.

node to the decision point plus from the decision point to the ingress node and back. If traffic descriptors are overestimated, the decision point terminates too little traffic so that several termination steps with sufficient inter-termination time are needed to remove SR-pre-congestion [5]. Thus, termination in the CL architecture also takes between one and two seconds depending on the network parameters and the overestimation of traffic descriptors. In contrast, RCFT does not depend on traffic descriptors so that their accuracy has no impact on the time to remove SR-pre-congestion.

5.4 Impact of Flow Rate Variability

We study the impact of flow rate variability. In our first experiment, we consider only homogeneous flows with a rate of 80 kbit/s. In our second experiment, we consider 80% flows with 20 kbit/s and 20% flows with 320 kbit/s, i.e., we do not modify the number of flows and their rate. We either change the packet size of the flows or the packet frequency to adapt the flow rates. Both alternatives lead to the same results as long as we use packet-size independent marking (PSIM) [23]. PSIM has been standardized in [12] and effects that packets are marked independently of their size.

Figure 3 shows the termination behavior for homogeneous and heterogeneous flow rates. The lines with the black markers are mean values over multiple simulation runs and show that the average termination behavior is the same for both traffic types. The lines with the white markers present the 10% and the 90% quantiles of these experiments. They show that the termination speed with RCFT is more variable for heterogeneous flows than for homogeneous flows. Moreover, more overtermination occurs but it is less than 10% in 90% of the cases. With homogeneous flows, overtermination is negligible.

5.5 Impact of Traffic Aggregation

We simulate bottleneck links with supportable rates of $SR \in \{1, 2, 4, 8\}$ Mbit/s and 100% SR-overload so that 12, 25, 50, or 100 flows can be supported. Figure 4 shows that RCFT works for all these aggregation levels and that the deviation from the mean of the time-dependent PCN rate diminishes with decreasing aggregation level. Here, we talk about the aggregation level on the bottleneck link and all flows may belong to different IEAs. In [5], we showed that measured rate termination with single marking leads to significant overtermination when the aggregation level is about 10 flows per IEA. The termination accuracy of RCFT does not suffer in that case.

5.6 Impact of Heterogeneous Termination Delays D_T

The termination delay D_T of flows passing a common bottleneck link may be different. We assume 50% flows with a small value of $D_T = 50$ ms and 50% flows with a large value of $D_T = 350$ ms. Figure 5 shows that the overall termination behavior is about the same as for flows with homogeneous termination delay of $D_T = 200$ ms. The traffic rate of the flows with shorter $D_T = 50$ ms diminishes faster than the one of the flows with longer $D_T = 350$ ms, but this does not affect the termination probabilities of the different flow types. The reason for this possibly counterintuitive result is the fact that all traffic experiences SR-pre-congestion for the same duration and, therefore, all flows have the same chance to be terminated.

5.7 Impact of Packet Loss

We consider extreme overload so that packet loss occurs. The bottleneck link has a supportable rate of 8

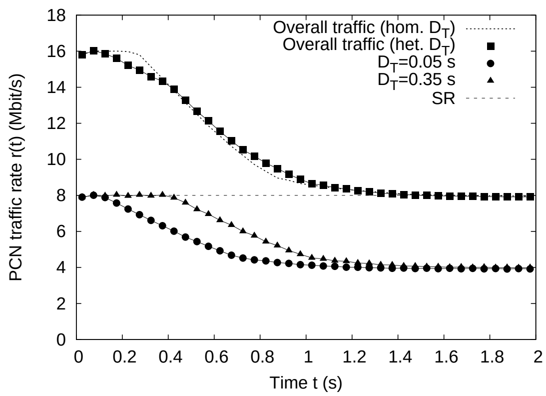


Fig. 5 Impact of heterogeneous termination delay D_T.

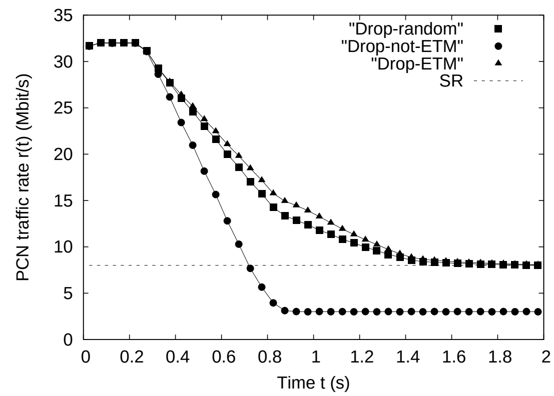
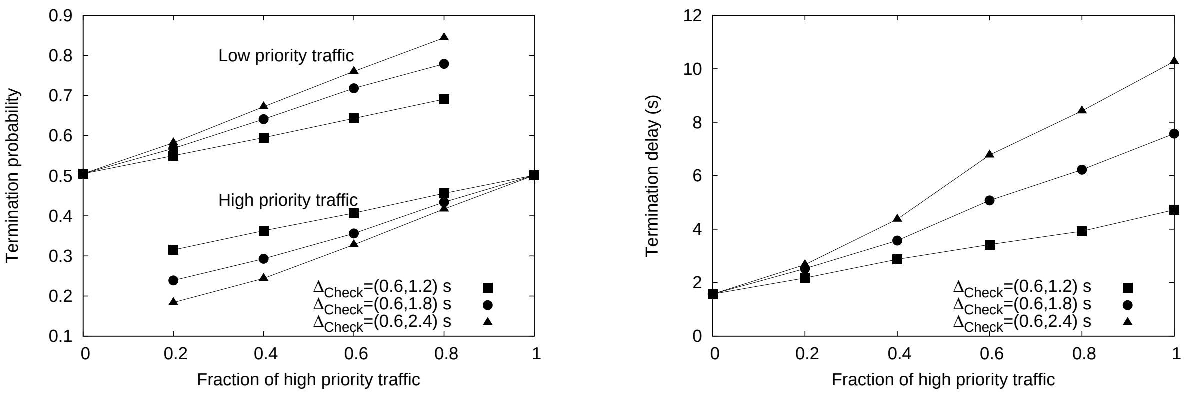


Fig. 6 Impact of packet loss.



(a) Termination probabilities.

(b) Time to remove SR-pre-congestion.

Fig. 7 Impact of flow termination prioritization using different Δ_{Check} values.

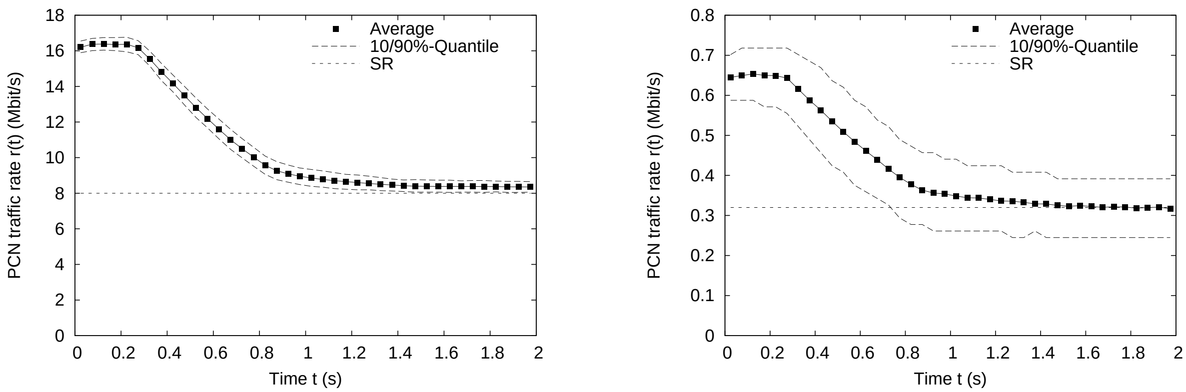
Mbit/s, a bandwidth of 16 Mbit/s, and during SR-pre-congestion 32 Mbit/s are offered to the link so that 50% packet loss occurs.

Figure 6 shows the termination behavior for three different packet drop policies. When packets that are not excess-traffic-marked are preferentially dropped (drop-not-ETM), then mainly excess-traffic-marked packets remain so that most of the flows are terminated at the first checkpoint after SR-pre-congestion occurs. This causes significant overtermination. When excess-traffic-marked packets are preferentially dropped (drop-ETM) or when packets are dropped independently of their PCN marking (drop-random), then overtermination is not observed. Thus, RCFT is perfectly compatible with the recommendations for dropping PCN traffic in case of overload in [12].

5.8 Termination Priorities Using Different Δ_{Check} Values

Some flows may be more important than others so that it is desirable to first terminate less important (low-priority) flows. In the CL architecture, the decision point chooses the flows for termination from a specific IEA and may preferentially select low-priority flows. Analogously, the egress node in our approach may terminate a low-priority flow instead of an excess-traffic-marked high-priority flow if at least one of the packets of the low-priority flow was recently excess-traffic-marked. However, these ideas work only if IEAs carry multiple flows.

RCFT offers an additional way to implement termination priorities. We configure the check interval of low-priority flows with the common value $\Delta_{Check} = 600$ ms and assign larger values of $\Delta_{Check} \in \{1.2, 1.8, 2.4\}$ s to high-priority traffic. Figure 7(a) shows the termination probabilities for different fractions of high-priority traffic. They are significantly lower for high-priority than for low-priority traffic. The difference between the termination probabilities increases with the duration of the check interval Δ_{Check} for high-priority flows. However, Figure 7(b) shows that also the time to fully remove SR-pre-congestion increases with an increasing duration of the check interval Δ_{Check} for high-priority



(a) $SR = 8$ Mbit/s, $n_{init} = 1205$ on/off flows ($\sim$ 16 Mbit/s). (b) $SR = 320$ kbit/s, $n_{init} = 48$ on/off flows ($\sim$ 640 kbit/s).

Fig. 8 Impact of on/off traffic and flow aggregation shortly after sudden SR-overload.

flows. This is problematic if the fraction of high-priority flows is large. Therefore, the value of Δ_{Check} for high-priority flows should be configured only moderately larger than for low-priority flows.

5.9 Impact of On/Off Traffic

Variable bitrate streams like compressed voice and video traffic present a special challenge for measurement- or feedback-based AC and FT systems as the rate of admitted traffic aggregates fluctuates over time. On/off sources like compressed voice, e.g. from the "internet Low Bitrate Codec" (iLBC) [44], are an extreme case of that category as they induce more variation than video traffic with the same number of flows. We first review a traffic model for the iLBC and then we use it to study PCN-based traffic control in the presence of on/off traffic.

5.9.1 Traffic Model for the iLBC Codec

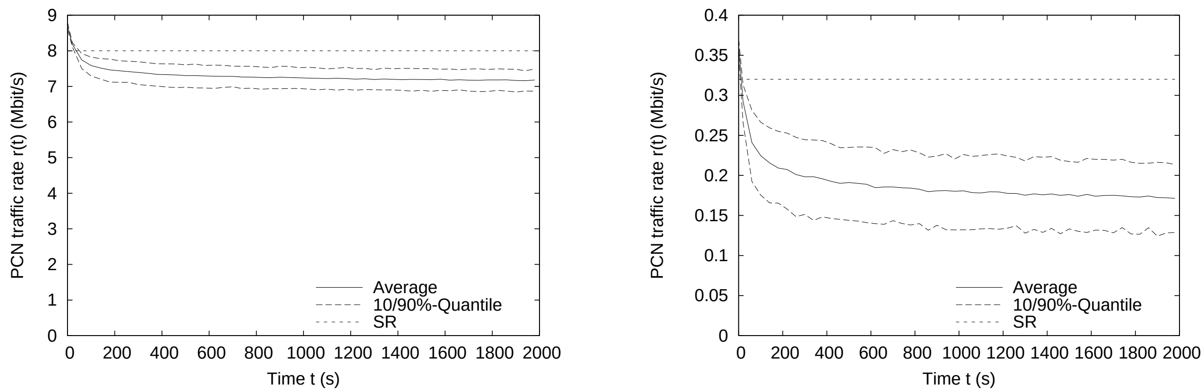
In [45] voice samples have been coded with the iLBC and the resulting traffic traces have been analyzed to provide a simulation model of iLBC traffic over IP networks. In the on-phase, packets with 102 bytes are sent every 30 ms while in the on-phase packet generation is suppressed. We denote the bitrate of a flow in the on-phase by $R_{on} = \frac{102 \text{ bytes}}{30 \text{ ms}} = 27.2$ kbit/s. The average duration of an on-phase is $E[D_{\text{on}}] = 11.00$ s, the average duration of an off-phase is $E[D_{\text{off}}] = 11.54$ s. This results in a voice activity factor of $\alpha = \frac{11.00 \text{ s}}{11.00 \text{ s}+11.54 \text{ s}} = 0.4880$ and an average bitrate of $R_{iLBC} = \frac{102 \text{ bytes}}{11.00 \text{ s}+11.54 \text{ s}} = 13.27$ kbit/s. The durations of the on- and off-phase can be modeled by exponential distributions of the form $P(D_{\{\text{on,off}\}} \leq t) = 1 - \exp(-\beta_{\{\text{on,off}\}} \cdot t)$ with $\beta_{\{\text{on,off}\}} = \frac{1}{E[D_{\{\text{on,off}\}}]}$.

5.9.2 Termination Behavior Shortly after Sudden Overload

In our first experiment, we consider a bottleneck link with $SR = 8$ Mbit/s. We start the simulation with 1205 iLBC-coded flows. The fraction α of them is in the on-phase (588) to generate almost 16 Mbit/s which is an SR-overload of 100%. Figure 8(a) illustrates the termination behavior of RCFT over the first two seconds. RCFT quickly terminates half of the traffic within 2 s so that SR-overload is removed. The 10%- and 90%-quantiles are about as tight as for CBR traffic (see Figure 3), i.e., the simulation behavior is about the same in any simulation run.

We perform the same experiment for $SR = 320$ kbit/s with only 48 iLBC-coded flows, α of them in their on-phase (23), so that the initial traffic load is almost 640 kbit/s. Figure 8(b) presents the simulated termination behavior. SR-overload is also quickly removed within 2 s. The 10% and 90% quantiles of the measured PCN traffic rate are relatively farer away from the average values than in Figure 8(a). That means, in some simulation runs the traffic rate is reduced faster or more slowly. In particular, the 10% quantiles exhibit some overtermination in a few simulation runs. This happens if a majority of flows turns into the on-phase so that they produce lots of traffic and many flows are terminated. When a majority of the remaining flows turns into the off-phase again, the resulting PCN traffic rate is clearly lower than SR. This phenomenon is visible only for low flow aggregation.

Thus, RCFT works well with on/off traffic in the sense that SR-pre-congestion is quickly detected and removed. If the traffic bundle is small in terms of flows, slight overtermination may occur in some cases.



(a) $SR = 8$ Mbit/s, $n_{init} = 1205$ on/off flows ($\sim$ 16 Mbit/s).

(b) $SR = 320$ kbit/s, $n_{init} = 48$ on/off flows ($\sim$ 640 kbit/s).

Fig. 9 Impact of on/off traffic and flow aggregation in the long run after sudden SR-overload; the initial SR-pre-congestion is not visible as most of it is removed within less than 1 s.

5.9.3 Termination Behavior in the Long Run after Sudden Overload

We perform again the same experiments but the simulation runs are 1000 times longer. The termination behavior is displayed in Figures 9(a) and 9(b) for $SR = 8$ Mbit/s and $SR = 320$ kbit/s. For $SR = 8$ Mbit/s the average PCN traffic rate is quickly reduced from 16 Mbit/s to slightly more than 7 Mbit/s and stays constant. Thus, visible but moderate overtermination occurs. The 10% and 90% quantiles are rather tight. We get different results for $SR = 320$ kbit/s. The average traffic rate is also quickly reduced from 640 kbit/s to $SR = 320$ kbit/s, but then still decreases continuously over time so that it falls down to almost half of SR after 2000 s. The 10% and 90% quantiles are rather wide, i.e., some simulation runs suffered even more from overtermination, others somewhat less. The reason for overtermination is the same as explained above. This experiment proves that RCFT does not work properly with on/off traffic in the presence of low flow aggregation because the observed large overtermination over time is not acceptable. However, this is not a specific phenomenon for RCFT. The reason rather lies in the very nature of on/off traffic so that any FT mechanism leads to similar results.

5.9.4 Termination of Admitted Traffic under Normal Operation

Admitted flows must not be terminated under normal operation without sudden overload that may arise from external factors, e.g., rerouted traffic. The results presented above raise the question whether PCN is able to cope with on/off traffic in the absence of sudden overload, at least for low flow aggregation. One solution to avoid or to mitigate the termination of admitted flows due to the statistical rate fluctuations of on/off traffic is the configuration of a large headroom between the admissible and supportable rates AR and SR. In the following, we evaluate whether a reasonably large headroom can effectively avoid the termination of admitted flows in the presence of on/off traffic.

In our first experiment, we consider a bottleneck link that is configured with an admissible rate of $AR = 4$ Mbit/s so that $n_{AR} = \lfloor \frac{AR}{R_{iLBC}} \rfloor = 301$ flows can be admitted on average. The average duration of a flow is $E[F] = 90$ s and its distribution is exponential. Admission requests arrive according to a Poisson process for which we choose a base arrival rate of

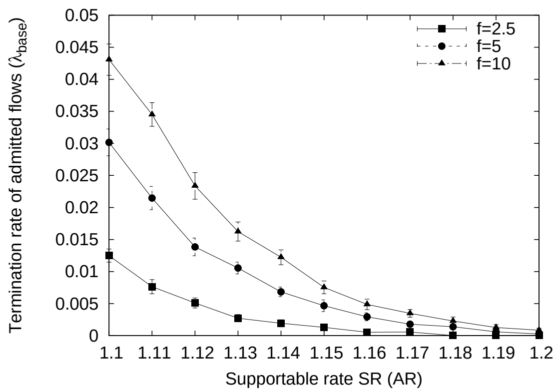
$$\lambda_{base} = \frac{n_{AR}}{E[F]} \quad (1)$$

which is $\lambda_{base} = 3.344 \frac{1}{\text{s}}$ in our specific example of $AR =$ 4 Mbit/s. To stress the system, we assume a flash crowd factor of $f_{crowd}^{flash} \in \{2.5, 5, 10\}$ like in [4]. The effective arrival rate of admission requests is then

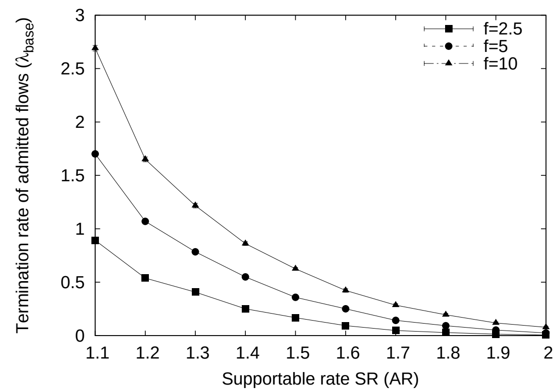
$$\lambda_{eff} = f_{crowd}^{flash} \cdot \lambda_{base} \quad (2)$$

so that the actual arrival rates are $\lambda_{eff} \in \{8.37, 16.75, 33.5\} \frac{1}{\text{s}}$. We perform probe-based AC with implicit probing as described in [4].

The simulation starts with an empty system which quickly admits so many flows that flow blocking starts. After 50 s, the number of termination events n_{term} are counted for a duration of $D_{sim} = 2000$ s to calculate the termination rate $\tau = \frac{n_{term}}{D_{sim}}$. Sufficient simulation runs are performed to produce small confidence intervals for the termination rate τ with a confidence level of 95%. Figure 10(a) presents the measured termination rates for $AR = 4$ Mbit/s. The x-axis shows the configured supportable rate SR (relative to the given AR) and the



(a) $AR = 4$ Mbit/s.



(b) $AR = 160$ kbit/s.

Fig. 10 Simulated termination rates.

y-axis shows the termination rate as a fraction of the base arrival rate λ_{base}. The rate λ_{base} can be imagined as the rate of admissible flows (with rate R_{iLBC}) in the absence of on/off traffic.

The termination rates clearly decrease with increasing supportable rate since more headroom can better avoid termination in case of rate fluctuations due to on/off traffic. The termination rates increase almost linearly with the flash crowd factor f_{crowd}^{flash} because the number of admission requests that arrive when the PCN traffic rate is below AR scales with f_{crowd}^{flash}. For $AR = 4$ Mbit/s, the termination rates are rather low. They are in the order of 1% of the admissible flow rate for 10% headroom and quickly decrease to significantly lower termination rates in the order of 0.01% of the admissible flow rate for a headroom of 20%. Thus, flow termination due to on/off traffic can effectively be avoided with rather little headroom.

We now consider a bottleneck link with $AR = 160$ kbit/s, so that only $n_{AR} = 12$ flows can be carried on average for which a base arrival rate of $0.133\frac{1}{\mathrm{s}}$ is needed. The flash crowd factors increase them to effective arrival rates of $\lambda_{eff} \in \{0.33, 0.67, 1.33\}\frac{1}{\mathrm{s}}$.

We perform the same experiment like above. The simulated data are compiled in Figure 10(b). They are qualitatively similar to those in Figure 10(a). However, for only 10% headroom, the terminated flow rates are between 105% and 271% of the admissible flow rate in the absence of on/off traffic, depending on the flash crowd factor f_{crowd}^{flash}. With 100% headroom this fraction can be reduced only to about 5% which is not a tolerable fraction, yet. Thus, lots of headroom is needed to effectively reduce the termination rates in the presence of on/off traffic and low flow aggregation.

5.9.5 Analysis of the Termination Rate Using a Continuous-Time Markov Chain

The simulation results presented above show that it is possible to support on/off traffic with PCN-based flow control. However, many simulation runs are needed to achieve small confidence intervals, especially for the simulation of low termination rates. Therefore, we provide a continuous-time Markov chain (CTMC) based on which approximative termination rates can be derived.

Model. We can construct a CTMC based on our simulation model because inter-arrival times of admission requests, flow durations, as well as on-phases and off-phases of the flows follow exponential distributions. In our analysis the admissible and supportable rates are expressed in terms of active flows, i.e., $AR_d = \frac{AR}{R_{on}}$ and $SR_d = \frac{SR}{R_{on}}$. The model has a two-dimensional state space (i, j) where $i \geq 0$ indicates the number of admitted flows and $0 \leq j \leq SR_d$ is the number of flows in the on-phase. Thus, for valid tuples (i, j) holds $i \geq j$. To facilitate computations, we prune the infinite state space (i, j) to $i \leq i_{max}$ with $SR_d \ll i_{max}$. We explain the various types of state transitions in the model.

- There are state transitions when flows change from on-phase to off-phase and vice-versa. We model them by the equations

$$(i, j) \xrightarrow{j \cdot \beta_{on}} (i, j-1) \text{ for } 0 < j \leq SR_d \text{ and} \quad (3)$$

$$(i, j) \xrightarrow{(i-j) \cdot \beta_{off}} (i, j+1) \text{ for } 0 \leq j < SR_d \quad (4)$$

  with state transition rate β_{on} and β_{off}, respectively.
- In the presence of $j = SR_d$ active flows, the transition of an inactive flow to its on-phase triggers the termination of an active flow so that only the num-

ber of overall flows is decreased. Thus, we get

$$(i, SR_d) \xrightarrow{(i-SR_d)\cdot\beta_{off}} (i-1, SR_d) \text{ for } SR_d < i. \quad (5)$$

- The admission of new flows is modelled by the transitions

$$(i,j) \xrightarrow{\alpha\cdot\lambda_{eff}} (i+1, j+1) \text{ for } j \leq AR_d \text{ and} \quad (6)$$

$$(i,j) \xrightarrow{(1-\alpha)\cdot\lambda_{eff}} (i+1, j) \text{ for } j \leq AR_d. \quad (7)$$

They imply that flows start after admission with probability α in their on-phase and with probability $(1-\alpha)$ in their off-phase.

- The normal completion of active and inactive flows is described by

$$(i,j) \xrightarrow{j\cdot\mu} (i-1, j-1) \text{ and} \quad (8)$$

$$(i,j) \xrightarrow{(i-j)\cdot\mu} (i-1, j) \quad (9)$$

with $\mu = \frac{1}{E[F]}$.

A snapshot of a state transition diagram is provided in Figure 11 to visualize the CTMC. The state transition rates are compiled in the state matrix $\mathbf{Q}$ relative to a linearization of the two-dimensional state space. The stationary state distribution x_s of this CTMC is a row vector and can be calculated by solving the equation system

$$x_s \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 1 \text{ and} \quad (10)$$

$$x_s \cdot \mathbf{Q} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad (11)$$

The probability for i admitted flows can be derived as

$$p(i) = \sum_{0\leq j\leq i} x_s(i,j). \quad (12)$$

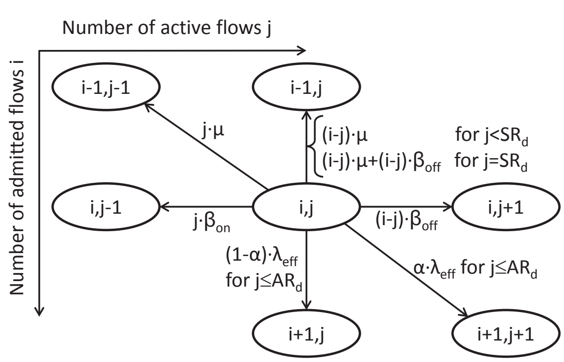


Fig. 11 A snapshot of the state transition diagram visualizes the CTMC.

The probability for j active flows can be calculated as

$$p(j) = \sum_{j\leq i\leq i_{max}} x_s(i,j). \quad (13)$$

The termination rate can be computed by

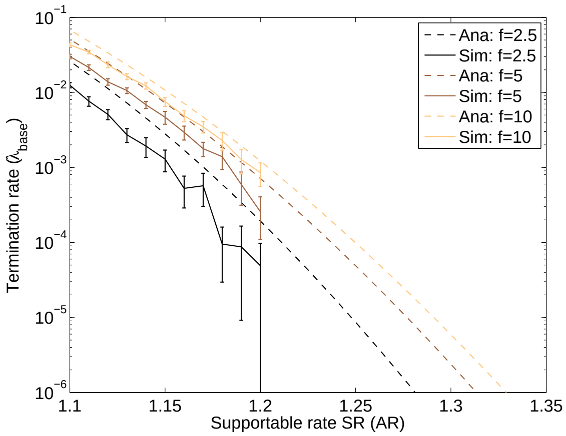
$$\tau = \sum_{SR_d\leq i\leq i_{max}} x_s(i, SR_d)\cdot(i-SR_d)\cdot\beta_{off} \quad (14)$$

Evaluation. Figures 12(a) and 12(b) compare the simulated and analytically calculated termination rates. The analytical values are in the correct order of magnitude but overestimate the expected results. The simulation implements the timing of the system correctly while the analytical model assumes immediate actions as soon as flow states change, which probably causes the deviation. An example: the analytical model assumes that a flow is immediately terminated as soon as more than SR_d flows become active. However, excess traffic marking detects SR-overload by the depletion of the underlying token bucket. This depletion requires some time during which another flow may complete or switch to its off-phase. In Figure 12(b), the analytical curves do not start at a supportable rate of 1.1 because the smallest configurable headroom for $AR_d = 6$ is $SR_d = 7$ in the analysis which leads to $SR = \frac{7}{6}\cdot AR_d = 1.17\cdot AR_d$.

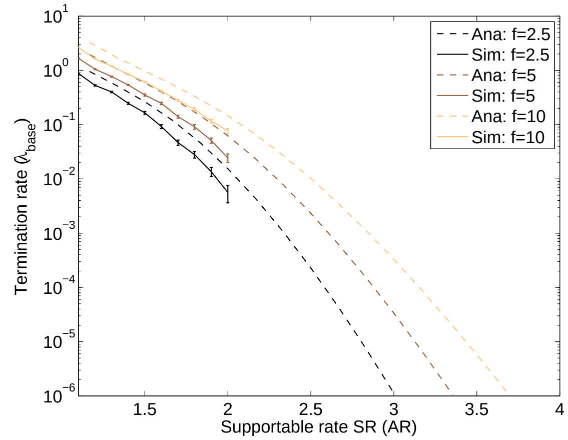
In spite of the deviation of the analytical data from the simulation results, the model is useful to quickly give upper bounds for configurations with very large headroom. Such headroom leads to very small termination rates that require lots of simulation effort for accurate results. Therefore, we stopped the simulation of termination rates in Figures 12(a) and 12(b) at supportable rates of $1.2\cdot SR$ and $2\cdot SR$, respectively.

With an admissible rate of $AR = 4$ Mbit/s, $AR_d = \frac{AR}{R_{on}} = \frac{4\text{ Mbit/s}}{27.2\text{ kbit/s}} = 147.05 \approx 147$ active flows can be supported without causing AR-pre-congestion. Figure 12(a) suggests that with a headroom of 35% ($SR_d =$ 199 flows) the flow termination rate is only $10^{-6}\cdot\lambda_{base}$, i.e., only one out of 10^6 admitted flows is terminated, which seems acceptable. With $AR = 160$ kbit/s we get $AR_d = \frac{160\text{ kbit/s}}{27.2\text{ kbit/s}} = 5.88 \approx 6$ flows. In that experiment a headroom of about 250% ($SR_d = 21$ flows) is needed to achieve the same goals. The use of PCN-based flow control is not economic under these conditions. When a simple counting-based admission control admits at most $\frac{AR}{R_{iLBC}}$ flows, then AR bandwidth is used on average. The admitted flows can generate a maximum bitrate of $\frac{AR}{R_{iLBC}}\cdot R_{on} = \frac{AR}{\alpha} = 2.05\cdot AR$ so that only a headroom of 105% is needed.

For a better understanding of this phenomenon, Figures 13(a) and 13(b) illustrate the complementary cumulative distribution function (CCDF) of admitted and

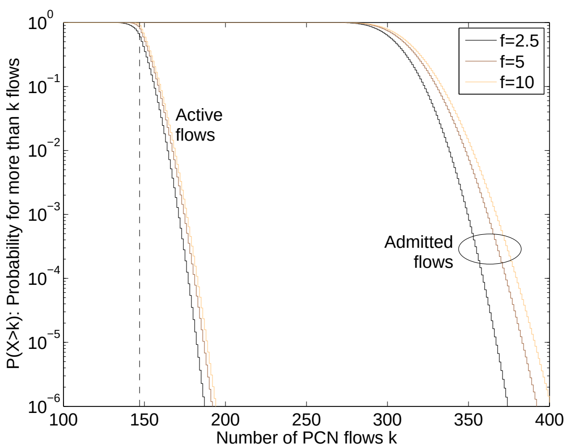


(a) $AR = 4$ Mbit/s.

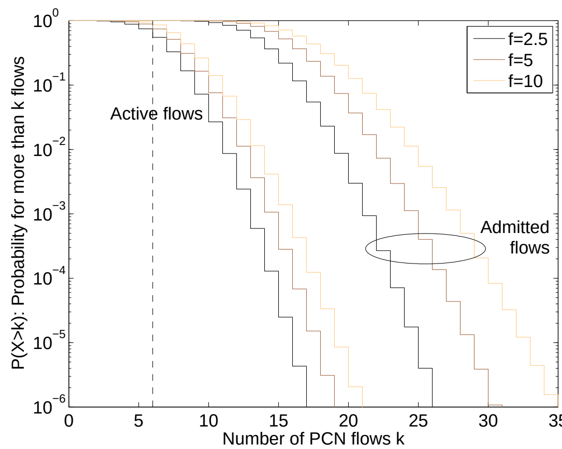


(b) $AR = 160$ kbit/s.

Fig. 12 Simulated and analytically calculated termination rates.



(a) $AR = 4$ Mbit/s, $SR = 8$ Mbit/s.



(b) $AR = 160$ kbit/s, $SR = 640$ kbit/s.

Fig. 13 CCDF for the number of admitted and active PCN flows on bottleneck links with different aggregation levels.

active flows. The admissible rates are $AR_d = 147$ and $AR_d = 6$ active flows, respectively, and the headroom is configured with 100% and 300% in order to make flow termination an extremely unlikely event. We observe in the figures that under these conditions up to 400 and 35 flows are admitted, respectively. Out of them up to 195 and 21 are simultaneously active. To avoid flow termination for them, the supportable rate needs to be 33% and 250% larger than the corresponding admissible rates in terms of active flows (147 and 6). Thus, PCN-based flow termination is efficient even in the presence of on/off traffic provided that flow aggregation is large enough.

6 Conclusion

We have proposed regular-check based flow termination (RCFT) as a simple method for flow termination using pre-congestion notification (PCN). Together with marked-signalling-based admission control it allows for simple admission control and flow termination in Differentiated Services IP networks. It is clearly simpler than the currently proposed experimental specification [3] for multiple reasons that we have highlighted in this paper.

We have investigated RCFT under various challenging networking conditions. RCFT requires the duration of the check interval as single configuration parameter. We provided a simple configuration rule for which RCFT quickly terminates SR-overload without overtermination. Moreover, it is more robust than other marked flow termination methods that have been proposed before. However, overtermination may occur in the presence of on/off traffic over time. This is not a special finding for RCFT, it is rather a typical behavior of systems with PCN-based admission control and flow termination. We provided an analytical model and showed that the termination probability for admitted flows can be kept very small with moderate headroom (in terms of supportable rate) provided that the aggregation level of flows is large enough. The analysis may be used for the configuration of general PCN-based flow control. Further work may elaborate our study of on/off traffic for multiplexed video traffic which requires more complex simulation and analytical models.

Acknowledgments

The authors would like to thank Joe Babiarz, Bob Briscoe, Kwok Chan, Anna Charny, Ruediger Geib, and Phil Eardley for fruitful discussions. This work was funded by Nortel Networks, Ottawa, and Deutsche Forschungsgemeinschaft (DFG) under grants TR257/18-3 and MM2727/1-1. The authors alone are responsible for the content of the paper.

Appendix

A list of abbreviations is given in Table 1.

Table 1 List of frequently used acronyms.

Acronym	Meaning
Δ_{Check}	duration of the check interval
AC	admission control
AR	admissible rate
CL	"Controlled Load" PCN architecture [3]
CLE	congestion level estimate
DS	Differentiated Services
D_T	termination delay
EMR	rate of ETM-traffic measured by the egress node
ETM	excess-traffic marked
FT	flow termination
IEA	ingress-egress aggregate
IETF	Internet Engineering Task Force
IR	rate of PCN traffic sent and measured by the ingress node
MFT	marked flow termination
MFT-IEA	MFT for IEAs
MFT-IF	MFT for individual flows
MPLS	Multiprotocol Label Switching
MRT	measured rate termination
MSAC	marked-signalling based admission control
NM	not-marked
NMR	rate of NM-traffic measured by the egress node
QoS	Quality of Service
PCN	pre-congestion notification
PSIM	packet size independent marking
PSDM	packet-specific dual marking
RACF	Resource and Admission Control Function
RCFT	regular-check based flow termination
RSVP	Resource reSerVation Protocol [1]
SIP	Session Initiation Protocol
SR	supportable rate
ThM	threshold-marked
TMR	rate of TM-traffic measured by the egress node
TR	termination rate

References

1. Braden, B., Zhang, L., Berson, S., Herzog, S., Jamin, S.: RFC2205: Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification (1997)
2. Wroclawski, J.: RFC2211: Specification of the Controlled-Load Network Element Service (1997)
3. Charny, A., Huang, F., Karagiannis, G., Menth, M., Taylor, T.: RFC 6661: PCN Boundary Node Behaviour for the Controlled Load (CL) Mode of Operation (2012)
4. Menth, M., Lehrieder, F.: Performance of PCN-Based Admission Control under Challenging Conditions. IEEE/ ACM Transactions on Networking **20** (2012)
5. Menth, M., Lehrieder, F.: PCN-Based Measured Rate Termination. Computer Networks **54** (2010) 2099 – 2116
6. Karagiannis, G., Bhargava, A.: Generic Aggregation of Resource ReSerVation Protocol (RSVP) for IPv4 and IPv6 Reservations over PCN Domains. http://tools.ietf.org/html/draft-ietf-tsvwg-rsvp-pcn (2012)
7. Menth, M., Geib, R.: Admission Control Using PCN-Marked Signaling. http://tools.ietf.org/html/draft-menth-pcn-marked-signaling-ac (2011)
8. Eardley (Ed.), P.: RFC5559: Pre-Congestion Notification (PCN) Architecture (2009)
9. Briscoe, B., Moncaster, T., Menth, M.: RFC 6660: Encoding 3 PCN-States in the IP header using a single DSCP (2012)
10. : ITU-T Recommendation Y.2111 (2006), Resource and Admission Control Functions in Next Generation Networks (2006)
11. Menth, M., Briscoe, B., Tsou, T.: Pre-Congestion Notification (PCN) – New QoS Support for Differentiated Services IP Networks. IEEE Communications Magazine **50** (2012)
12. Eardley (Ed.), P.: RFC5670: Metering and Marking Behaviour of PCN Nodes (2009)
13. Kelly, F., Key, P., Zachary, S.: Distributed Admission Control. IEEE Journal on Selected Areas in Communications **18** (2000) 2617–2628
14. Karsten, M., Schmitt, J.: Packet Marking for Integrated Load Control. In: IFIP/IEEE Symposium on Integrated Management (IM). (2005)
15. Menth, M., Lehrieder, F., Briscoe, B., Eardley, P., Moncaster, T., Babiarz, J., Charny, A., Zhang, X.J., Taylor, T., Chan, K.H., Satoh, D., Geib, R., Karagiannis, G.: A Survey of PCN-Based Admission Control and Flow Termination. IEEE Communications Surveys & Tutorials **12** (2010)
16. Zhang, X., Charny, A.: Performance Evaluation of Pre-Congestion Notification. In: International Workshop on

Quality of Service (IWQoS), Enschede, The Netherlands (2008)
17. Latre, S., De Vleeschauwer, B., Van de Meerssche, W., De Schepper, K., Hublet, C., Van Leekwijck, W., De Turck, F.: PCN Based Admission Control for Autonomic Video Quality Differentiation: Design and Evaluation. Journal of Network and Systems Management **19** (2011) 32 – 57
18. Arumaithurai, M., Geib, R., Rex, R., Fu, X.: Pre-Congestion Notification-based Flow Management in MPLS-based DiffServ Networks. In: IEEE International Performance Computing and Communications Conference (IPCCC), Phoenix, AZ, USA (2009)
19. Geib, R., Azanon-Teruel, E., Donaire-Arroyo, S., Ferrandiz-Cancio, A., Ralli-Ucendo, C., Romero-Bueno, F.: Service Deployment Experience in Pre-Commercial IPv6 Networks. UPGRADE **6** (2005)
20. Menth, M., Babiarz, J., Eardley, P.: Pre-Congestion Notification Using Packet-Specific Dual Marking. In: International Workshop on the Network of the Future (Future-Net), Dresden, Germany (2009)
21. Briscoe, B.: RFC6040: Tunnelling of Explicit Congestion Notification (2010)
22. Satoh, D., Ueno, H.: Cause and Countermeasure of Overtermination for PCN-Based Flow Termination. In: IEEE Symposium on Computers and Communications (ISCC), Riccione, Italy (2010)
23. Menth, M., Lehrieder, F.: PCN-Based Marked Flow Termination. Computer Communications **34** (2011) 2082 – 2093
24. Lehrieder, F., Menth, M.: Marking Conversion for Pre-Congestion Notification. In: IEEE International Conference on Communications (ICC), Dresden, Germany (2009)
25. Charny, A., Zhang, J., Karagiannis, G., Menth, M., Taylor, T.: RFC 6662: PCN Boundary Node Behaviour for the Single Marking (SM) Mode of Operation (2012)
26. Lehrieder, F., Menth, M.: PCN-Based Flow Termination with Multiple Bottleneck Links. In: IEEE International Conference on Communications (ICC), Dresden, Germany (2009)
27. Menth, M., Hartmann, M.: Threshold Configuration and Routing Optimization for PCN-Based Resilient Admission Control. Computer Networks **53** (2009) 1771 – 1783
28. International Telecommunications Union: Clause 3 – Multi-Level Precedence and Preemption (MLPP). ITU Recommendation Q.955.3 (1994)
29. International Telecommunications Union: Integrated Services Digital Network (ISDN) - General Structure and Service Capabilities - Multi-Level Precedence and Preemption. ITU Recommendation I.255.3 (1990)
30. Schulzrinne, H., Polk, J.: RFC 4412: Communications Resource Priority for the Session Initiation Protocol (SIP) (2006)
31. 3GPP: TS 22.067: enhanced Multi-Level Precedence and Pre-emption service (eMLPP); Stage 1 (2009)
32. 3GPP: TS 23.067: enhanced Multi-Level Precedence and Pre-emption service (eMLPP); Stage 2 (2009)
33. 3GPP: TS 24.067: enhanced Multi-Level Precedence and Pre-emption service (eMLPP); Stage 3 (2009)
34. Kingston, J.: Dynamic Precedence for Military IP Networks. In: Military Communications Conference (MILCOM). (2000)
35. McCann, C., Elmasry, G., Russell, B., Welsh, B.: A Measurement-Based Approach for Multilevel Admission of Heterogeneous Traffic in Wireless Ad-Hoc Networks. In: Military Communications Conference (MILCOM), Taunton, MA, USA (2004)
36. Fineberg, V.: Specification of the Military Precedence and Preemption in the DS-TE Networks. In: Military Communications Conference (MILCOM), DISA, Falls Church, VA, USA (2004)
37. Shan, T., Yang, O.W.W.: Bandwidth Management for Supporting Differentiated-Service-Aware Traffic Engineering. IEEE Transactions on Parallel and Distributed Systems **18** (2007) 1320 – 1331
38. de Oliveira, J.C., Scoglio, C., Akyildiz, I.F., Uhl, G.: New Preemption Policies for DiffServ-Aware Traffic Engineering to Minimize Rerouting in MPLS Networks. IEEE/ACM Transactions on Networking **12** (2004) 733 – 745
39. Chaieb, I., Le Roux, J.L., Cousin, B.: A New Pre-emption Policy For MPLS-TE Networks. In: IEEE International Conference on Networks (ICON). (2007)
40. Goldberg, J.B., Dasgupta, S., de Oliveira, J.C.: Bandwidth Constraint Models: A performance study with preemption on link failures. In: IEEE Globecom, San Francisco, CA, USA (2006)
41. Bar-Noy, A., Canetti, R., Kutten, S., Mansour, Y., Schieber, B.: Bandwidth Allocation with Preemption. SIAM Journal on Computing **28** (1999) 1806 – 1828
42. Garay, J.A., Gopal, I.S.: Call Preemption in Communication Networks. In: IEEE Infocom. (1992)
43. Dogar, F.R., Aslam, L., Uzmi, Z.A., Abbasi, S., Kim, Y.C.: Connection Preemption in Multi-Class Networks. In: IEEE Globecom, San Francisco, CA, USA (2006)
44. Andersen, S., Duric, A., Astrom, H., Hagen, R., Kleijn, W., Linden, J.: RFC3951: Internet Low Bit Rate Codec (iLBC) (2004)
45. Menth, M., Binzenhöfer, A., Mühleck, S.: Source Models for Speech Traffic Revisited. IEEE/ACM Transactions on Networking **17** (2009) 1042–1051

Frank Lehrieder studied computer science and business administration at the University of Würzburg/Germany and University Antonio de Nebrija in Madrid/Spain. In May 2008, he received his Master degree in computer science from University of Würzburg. Now, he is a researcher at the Chair of Communication Networks in Würzburg and pursuing his PhD. His special interests are performance evaluation and resource management for peer-to-peer networks and Future Internet infrastructures.

Michael Menth is professor at the Department of Computer Science at the University of Tuebingen/Germany and chairholder of Communication Networks. He received a diploma and a PhD degree in 1998 and 2004 from the University of Wuerzburg/Germany. Prior he was studying computer science at the University of Texas at Austin and worked at the University of Ulm/Germany. His special interests are performance analysis and optimization of communication networks, resource management, resilience issues, smart grids, and future Internet. He holds numerous patents and received various scientific awards for innovative work.