PCN-Based Measured Rate Termination☆

Michael Menth, Frank Lehrieder

University of Würzburg, Departement of Computer Science, Germany

Abstract

Overload in a packet-based network can be prevented by admitting or blocking new flows depending on its load conditions. However, overload can occur in spite of admission control due to unforseen events, e.g., when admitted traffic is rerouted in the network after a failure. To restore quality of service for the majority of admitted flows in such cases, flow termination has been proposed as a novel control function. We present several flow termination algorithms that measure so-called pre-congestion notification (PCN) feedback. We analyze their advantages and shortcomings in particular under challenging conditions. The results improve the understanding of PCN technology which is currently being standardized by the Internet Engineering Task Force (IETF).

Keywords: Flow termination, admission control, resilience, QoS, Differentiated Services

1. Introduction

DiffServ networks [1] offer preferred treatment of high-priority traffic so that premium traffic like voice or video do not suffer packet loss or delay caused by other traffic which is carried over the same transmission links. However, if the rate of prioritized traffic is too large, overload of high-priority traffic may occur and lead to extensive packet loss and delay for prioritized traffic, too. This can happen since normal DiffServ networks lack an admission control (AC) function which admits high-priority flows to the network only if sufficient free capacity is still available for this traffic class.

The Internet Engineering Task Force (IETF) currently standardizes pre-congestion notification (PCN) [2]. PCN gives warnings to egress nodes nodes of a DiffServ domain [1] if the load of high-priority traffic has exceeded a critical level on some link. This information is used to implement a lightweight AC in the sense that per-flow states need to be kept only where flows enter and leave the domain.

Under normal conditions, PCN-based AC can enforce quality of service (QoS) in DiffServ networks. However, overload can occur in spite of AC due to unforseen events. For instance, admitted PCN traffic may be rerouted in case of a network failure and cause overload on backup links, or the rate of multiple admitted PCN flows may suddenly increase. To restore then a "controlled load" situation [3], flow termination (FT) has been proposed in the PCN context as an additional flow control function.

☆This work is funded by Nortel Networks, Ottawa, and Deutsche Forschungsgemeinschaft (DFG) under grant TR257/18-2. The authors alone are responsible for the content of the paper.

*Corresponding author.

Email address: menth@informatik.uni-wuerzburg.de (Michael Menth)

In [4] we have presented a survey of PCN-based AC and FT. In this paper, we investigate the performance of FT methods that rely on measured PCN feedback (measured rate termination, MRT). We show that some of them terminate more traffic than desired under certain conditions while others take quite a while to remove excess traffic. In addition, we propose countermeasures that improve the performance. This study covers in particular the FT algorithms that are eventually standardized. Our analytical and simulation results explain why these algorithms were chosen and reveal which conditions need to be met for proper operation.

The paper is structured as follows. Section 2 explains PCN, metering and marking algorithms as well as various FT algorithms. Section 3 reviews related work. Section 4 studies MRT methods under challenging conditions. Finally, Section 5 summarizes our findings and Section 6 draws conclusions. The appendix contains a list of frequently used acronyms.

2. Flow Termination Based on Pre-Congestion Notification (PCN)

In this section we explain the general idea of PCN-based admission control (AC) and flow termination (FT) and illustrate their application in a DiffServ domain in the Internet. We explain the metering and marking algorithms briefly and the FT algorithms in more detail.

2.1. Pre-Congestion Notification (PCN)

PCN defines a new traffic class for DiffServ networks that receives preferred forwarding treatment. Moreover, PCN provides feedback information from inside a DiffServ domain for AC and FT decisions at the borders to support QoS. To that end, PCN introduces an admissible and a supportable rate threshold

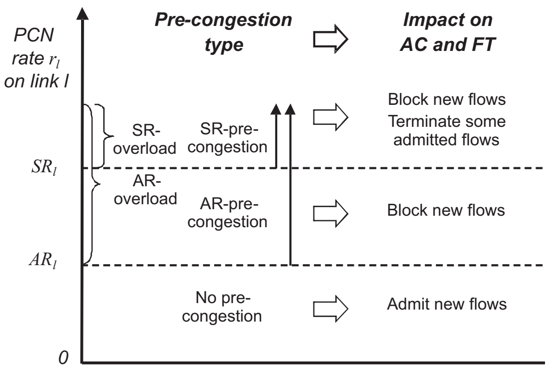


Figure 1: The admissible and the supportable rate (AR_l, SR_l) define three types of pre-congestion on link l.

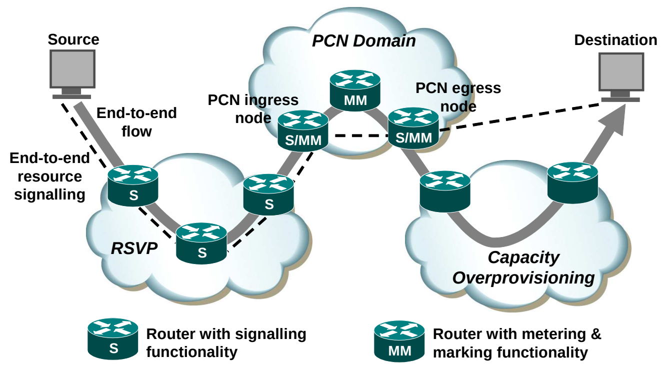


Figure 2: PCN-based AC is triggered by admission requests from external signalling protocols and guarantees QoS within a single PCN domain.

(AR_l, SR_l) for each link l of the DiffServ domain. This implies three different load regimes 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 the rate above AR_l is AR-overload. In this state, no further PCN flows should be admitted that would be carried over this link. If the PCN traffic rate r_l is above SR_l, the link is SR-pre-congested and the rate above SR_l is SR-overload. In this state, some already admitted flows that are carried over this link should be terminated to reduce the PCN rate r_l below SR_l.

2.2. Application of PCN in the Internet

PCN-based flow control assumes that some end-to-end signalling protocol (e.g. RSVP or SIP) or a similar mechanism requests admission for a new flow to cross a so-called PCN domain which is similar to the IntServ-over-DiffServ concept [5]. Thus, PCN-based AC and FT are per-domain QoS mechanisms and present an alternative to RSVP clouds or extreme capacity overprovisioning. This is illustrated in Figure 2. Traffic enters a PCN domain only through PCN ingress nodes and leaves it only through PCN egress nodes. Ingress nodes set a special header codepoint to make the packets distinguishable from other traffic and the egress nodes clear the codepoint. The nodes within a PCN domain are PCN nodes. They monitor the PCN traffic rate on their links and possibly re-mark the traffic in case of AR- or SR-pre-congestion. PCN egress nodes evaluate the markings of the traffic and send the results to the AC and FT entities of the PCN domain. In the following, we assume for simplicity reasons that the AC and FT entities are collocated with the ingress nodes of the traffic. Centralized AC and FT entities are also discussed for which the findings of this study are also valid.

2.3. PCN Metering and Marking

When entering the PCN domain, all PCN packets are marked with “not-marked” (NM). PCN nodes re-mark PCN packets depending on the load regime using the algorithms presented in this section. Egress nodes evaluate the packet markings and report the results to the appropriate ingress nodes. The ingress nodes use this information to admit or block new admission requests or to terminate already admitted flows. We first describe the metering and marking algorithms in the context of PCN and then we explain marking models.

2.3.1. Algorithms

There are two basic marking strategies: threshold and excess traffic marking [6]. A token bucket based meter tracks whether a certain reference rate is exceeded. Threshold marking re-marks all packets as “threshold-marked” (TM) when the PCN traffic rate exceeds the reference rate. Its marking result clearly indicates whether the reference rate was exceeded or not, and it is useful for AC purposes. Excess marking re-marks only those packets as “excess-traffic-marked” (ETM) that exceed the reference rate. The rate of ETM-packets provides an estimate of the rate by which the reference rate was exceeded while the rate of non-ETM-packets corresponds to the reference rate. Excess traffic marking is especially useful for flow termination as it allows the estimation of the traffic rate to be terminated. Excess traffic marking can be implemented with only few modifications of existing hardware. Threshold marking is not difficult to implement, either, but requires more changes to existing implementations.

2.3.2. Marking Models

PCN can be deployed with dual and single marking. We explain them in the following.

Dual Marking. Dual marking uses both threshold and excess traffic marking per link in a PCN domain [7]. Threshold marking configured with the admissible rate as reference rate re-marks NM-packets to TM. Thus, all packets are re-marked to TM in case of AR-overload which gives a clear signal for AC decisions. In addition, excess traffic marking configured with the supportable rate as reference rate re-marks NM- or TM-packets to ETM. ETM-packets must never be re-marked to NM or TM. In case of SR-overload, exactly the SR-overload is marked with ETM which serves as a good rate estimate for flow termination unless ETM-packets are lost. For FT, NM- and TM-packets are equally treated and we denote them in the following also as non-ETM.

Single Marking. Single marking uses only excess traffic marking [8]. Its reference rate is set to the admissible rate and it re-marks NM-packets to ETM. Hence, an amount of traffic equivalent to AR-overload is ETM. AC should stop admission of further flows as soon as some ETM-packets arrive at the egress node. The supportable rates are related to the admissible rates and are calculated by

$$SR = u \cdot AR \quad (1)$$

where $u > 1$ is a network-wide unique and configurable parameter. In case of SR-pre-congestion, more than $\frac{u-1}{u}$ of the PCN traffic is ETM, and all ETM PCN traffic above that fraction should be terminated. The advantage of single marking compared to dual marking is that only two (NM, ETM) instead of three PCN codepoints (NM, TM, ETM) are needed for PCN marking which facilitates the encoding of PCN marks in IP headers. Furthermore, systems can be built almost with off-the-shelf components as excess traffic marking is already implemented in routers. However, dual marking solutions work more accurately than single marking solutions. This has been shown for AC in [9] and we will show it for FT in this study.

2.4. Algorithms for PCN-Based Flow Termination

We review measured rate termination (MRT) methods in detail and briefly describe the idea of marked packet termination (MPT) which is a non-preferred alternative for the implementation of PCN-based FT. We describe them for dual and single marking. We omit the description of PCN-based AC algorithms and refer the interested reader to [4].

2.4.1. Measured Rate Termination (MRT)

MRT requires the notion of an ingress-egress aggregate (IEA) which is the set of flows between a specific ingress and egress node. With MRT, the PCN egress node measures the rates of NM-, TM-, and ETM-traffic (NMR, TMR, EMR) per IEA based on intervals of duration D_{MI} and signals them as so-called PCN feedback to the corresponding ingress node. When the ingress node receives these measurement reports, it carries out the procedures explained in the following to perform FT. We review different MRT types which can be adapted to dual and single marking. All of them assume that ingress nodes know signalled maximum rates for admitted flows. They need them to configure policers so that only admitted PCN traffic can enter the PCN domain. Ingress nodes can also use this information to select appropriate sets of flows for termination.

MRT with Directly Measured Termination Rates (MRT-DTR). With MRT-DTR, the ingress node calculates per IEA an estimate of the termination rate TR that needs to be terminated. It chooses a set of flows with an overall rate of TR from the corresponding IEA and terminates them. With dual marking, the egress node takes the rate EMR of ETM-traffic as a direct estimate for TR. With single marking, TR is calculated by $TR = \max(0, NMR + EMR - u \cdot NMR) = \max(0, EMR - (u - 1) \cdot NMR)$.

MRT with Measured Sustainable Aggregate Rates (MRT-SAR). With MRT-SAR, the ingress node calculates an estimate of the sustainable aggregate rate (SAR) per IEA which is the traffic rate that can be carried without causing SR-pre-congestion. The ingress node chooses a set of flows with an overall rate of SAR from the corresponding IEA and terminates all other flows of the IEA. With dual marking, the rate of non-ETM-traffic ($NMR + TMR$) is taken as a direct estimate for SAR. With single marking, the sustainable aggregate rate is calculated by $SAR = u \cdot NMR$.

MRT with Indirectly Measured Termination Rates (MRT-ITR). With MRT-ITR, the ingress node first decides whether termination is required. In case of dual marking, this is indicated by $EMR > 0$ and in case of single marking, this is indicated by $u \cdot NMR < NMR + EMR$. If termination is required, the ingress node computes SAR values like in Section 2.4.1 and performs local rate measurement of the sent PCN traffic, the so-called sent PCN ingress rate (IR). Then, the termination rate is calculated by $TR = \max(0, IR - SAR)$ and a set of flows with a traffic rate equal to TR is chosen for termination.

2.4.2. Marked Packet Termination (MPT)

MPT works without rate measurement by ingress and egress nodes. Various proposals exist. For instance, the egress node maintains a credit counter for each admitted flow which is reduced by the amount of marked bytes received for that flow. When the counter becomes negative, the flow is terminated. Another version of MPT uses excess traffic marking with marking frequency reduction and terminates a flow as soon as one of its packets is ETM. These and other methods have been proposed in [10], their performance has been evaluated, and recommendations have been given for configuration.

3. Related Work

We first review related work regarding other marking mechanisms and stateless core concepts for AC because they can be viewed as historic roots of PCN. Then we give a short summary of related PCN studies.

3.1. Related Marking Mechanisms

We present RED and ECN because they can be seen as precursors of PCN marking.

3.1.1. Random Early Detection (RED)

RED was originally presented in [11], and in [12] it was recommended for deployment in the Internet. RED detects incipient congestion by measuring a time-dependent average buffer occupation avg in routers and randomly drops packets. The probability for packet drops increases with the measured buffer occupation avg. This is done to indicate congestion to TCP senders. The value of avg relates to the physical queue size which is unlike PCN metering that relates to the configured admissible or supportable rate.

3.1.2. Explicit Congestion Notification

Explicit congestion notification (ECN) is built on the idea of RED to signal incipient congestion to TCP senders in order to reduce their sending window [13]. Packets of non-ECN-capable flows can be differentiated by a "not-ECN-capable transport" (not-ECT, '00') codepoint from packets of a ECN-capable flow which have an "ECN-capable transport" (ECT) codepoint. In case of incipient congestion, RED gateways possibly drop not-ECT packets while they just switch the codepoint of ECT packets to "congestion experienced" (CE, '11') instead of discarding them. This improves the TCP throughput since retransmission of such packets is no longer needed. Both the ECN encoding in the packet header and the behavior of ECN-capable senders and receivers after the reception of a marked packet are defined in [13]. ECN comes with two different codepoints for ECT: ECT(0) ('10') and ECT(1) ('01'). They serve as nonces to detect cheating network equipment or receivers [14] that do not conform to the ECN semantics. The four codepoints are encoded in the (currently unused) bits of the Differentiated Services codepoint (DSCP) in the IP header which is a redefinition of the type of service octet [15]. The ECN bits can be redefined by other protocols and [16] gives guidelines for that. They are also reused for the encoding of PCN codepoints [17, 18, 19, 20].

3.2. Admission Control

We briefly review some AC methods that can be seen as forerunners of the PCN-based AC principle.

3.2.1. Admission Control Based on Reservation Tickets

To keep a reservation for a flow across a network alive, ingress routers send reservation tickets in regular intervals to the egress routers. Intermediate routers estimate the rate of the tickets and can thereby estimate the expected load. If a new reservation sends probe tickets, intermediate routers forward them to the egress router if they have still enough capacity to support the new flow and the egress router bounces them back to the ingress router indicating a successful reservation; otherwise, the intermediate routers discard the probe tickets and the reservation request is denied. The tickets can also be marked by a packet state. Several stateless core mechanisms work according to this idea [21, 22, 23].

3.2.2. Admission Control Based on Packet Marking

Gibbens and Kelly [24, 25, 26] theoretically investigated AC based on the feedback of marked packets whereby packets are marked by routers based on a virtual queue with configurable bandwidth. This core idea is adopted by PCN. Marking based on a virtual instead of a physical queue also allows to limit the utilization of the link bandwidth by premium traffic to arbitrary values between 0 and 100%. Karsten and Schmitt [27, 28] integrated these ideas into the IntServ framework and implemented a prototype. They point out that the marking can also be based on the CPU usage of the routers instead of the link utilization if this turns out to be the limiting resource for packet forwarding.

3.2.3. Resilient Admission Control

Resilient admission control admits only so much traffic that it still can be carried after rerouting in a protected failure scenario [29]. It is necessary since overload in wide area networks mostly occurs due to link failures and not due to increased user activity [30]. It can be implemented with PCN by setting the admissible rate thresholds AR_l low enough such that the PCN rate r_l on a link l is lower than the supportable rate threshold SR_l after rerouting.

3.3. Related Studies in PCN

An overview of PCN including a multitude of different PCN-based AC and FT mechanisms is given in [4]. Ramp marking is an implementation alternative to threshold marking. The impact of both marking schemes on packet marking probabilities has been investigated in [31]. It turned out that threshold marking is as good as ramp marking which excluded ramp marking from further consideration because it is more complex than threshold marking. A two-layer architecture for PCN-based AC and FT was presented in [32] and flow blocking probabilities have been studied for single aggregates and static load conditions. In [9], various AC methods have been studied under challenging conditions. The authors of [33] have investigated the applicability of PCN-based admission control for video services in access networks. [10] proposes various algorithms for PCN-based marked packet termination (MPT) and gives recommendations for their configuration. As they were proposed only for use with dual marking, they were adapted for use with single marking in [34] and their performance was evaluated. Overtermination due to multiple bottlenecks is investigated in [35]. [36] gives a high level summary about a large set of simulation results regarding PCN-based AC and FT and shows that these methods work well in most studied cases. In contrast to that work, we investigate in this paper especially those situations where PCN-based MRT does not work that well. We provide an understanding of these problems which helps to discern whether these methods are applicable in specific application scenarios. [37] evaluates the efficiency of resilient PCN-based AC with flow termination and other resilient AC methods without flow termination in optimally dimensioned networks. [38] studies how *AR* and *SR* thresholds should be set in PCN domains with resilience requirements and how link weights should be set in IP networks in order to maximize the admissible traffic rates. [39] investigates the impact of admissible and supportable rate thresholds on the admission and termination of on/off traffic.

4. Performance of Measured Rate Termination

In this section we study the three MRT methods MRT-DTR, MRT-SAR, and MRT-ITR with dual and single marking. We describe challenging conditions, investigate them by case-based analysis, mathematical analysis, or simulation, and present improvements.

4.1. *Impact of Overestimated Traffic Descriptors*

Traffic descriptors are usually communicated by end-to-end signalling protocols and used for the configuration of per-flow policers at ingress nodes. Therefore, they indicate rather an upper bound of expected flow rates than a reliable estimate of expected average flow rates. As they are the only information about rates of individual flows at the ingress nodes, they are used as rate estimates to choose flows for termination.

With MRT-DTR and MRT-ITR, flow termination chooses a set of PCN flows for termination such that their overall rates equal the termination rate TR. When traffic descriptors are larger than the actual flow rates, too little traffic is terminated so that undertermination occurs. As a consequence, another termination step is required.

With MRT-SAR, flow termination chooses a set of PCN flows whose overall rates equal the sustainable aggregate rate SAR and the set of all other flows is terminated. When traffic descriptors are larger than the actual flow rates, too little traffic remains after termination so that overtermination occurs. This is not acceptable and rules MRT-SAR out from further consideration.

4.2. *Impact of Biased Measurement Results*

The results of rate measurements are representative only if the measured rate is stable within a measurement interval. If it increases or decreases, the measurement results easily over- or underestimate the rate of the observed traffic at the end of the measurement interval. We identify two different sources for SR-overload on a link: (1) the PCN traffic rates of IEAs carried over the SR-pre-congested link have increased or (2) the number of IEAs carried over the SR-pre-congested link has increased. We further study these scenarios.

The first case may occur when multiple admitted flows sharing a common link synchronously start transmission or increase their traffic rates. Figure 3(a) shows the rate increase of the overall PCN traffic, the ETM-traffic, and the non-ETM-traffic of a particular IEA under these conditions. When traffic is terminated from the IEA, its rates of overall PCN traffic, the ETM-traffic, and non-ETM-traffic decrease like in Figure 3(b).

The second case may occur, e.g., when traffic from other IEAs is rerouted to a considered link which then becomes SR-pre-congested. Unlike in the first case, the overall PCN rate of IEAs sharing this link may stay the same while the SR-overload on the considered link increases. However, the rates of ETM-traffic of the involved IEAs increase while the rates of non-ETM-traffic decrease. This is illustrated in Figure 3(c). Figure 3(d) shows the development of ETM- and non-ETM-traffic rates of an IEA when SR-overload is removed without terminating flows of this particular IEA. This can happen when SR-pre-congestion is removed, e.g., by backup traffic flapping back to its primary path or by the termination of flows belonging other IEAs.

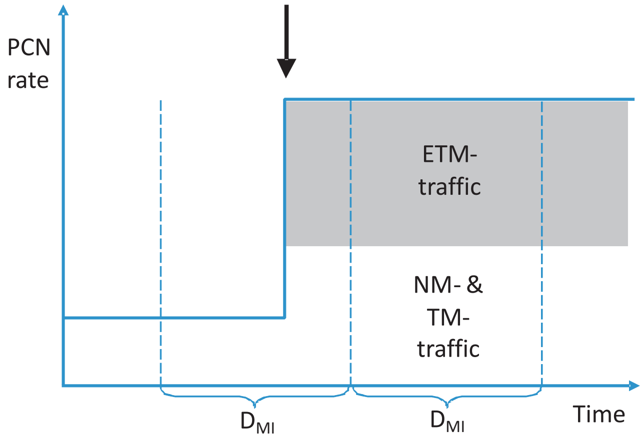
The presented changes of differently marked PCN traffic rates of an IEA may be observed during the measurement intervals of NMR, TMR, and EMR at the egress node and during the measurement intervals of the sent PCN ingress rate IR at the ingress node. They lead to biased measurement results which may cause over- or undertermination. In the following we discuss this for MRT-DTR and MRT-ITR with dual and single marking. We do not derive quantitative results as our intention is only to point out what can go wrong if mechanisms are not well designed and to present potential solutions if possible.

4.2.1. *Analysis of MRT-DTR*

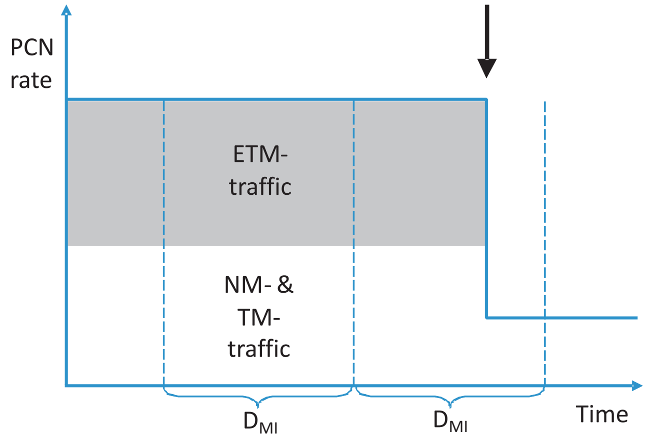
We consider MRT-DTR with dual and single marking when the rate of ETM-traffic increases like in Figures 3(a) and 3(c). The egress node's first measurement report covering ETM-packets is most likely to underestimate the rate of ETM-traffic. It is sent to the ingress node which uses it as an estimate for the termination rate TR. As a result, the first termination step results in undertermination and another termination step is needed.

We propose two different improvements. First, the ingress node should wait for a second PCN feedback indicating SR-pre-congestion because this is likely to capture the full SR-overload so that sufficient traffic can be terminated at once. This method works well for single and dual marking but introduces D_{MI} additional delay until termination starts. Second, the egress node may restart the measurement interval when a first ETM-packet arrives so that the rates are measured only during SR-pre-congestion. Then, the first measurement report is likely to reflect the full SR-overload so that the ingress node can terminate enough traffic in one shot. For single marking the arrival of ETM-packets at the egress node can be a sign for AR- or for SR-pre-congestion so that the restart of the measurement interval with beginning SR-pre-congestion cannot be enforced.

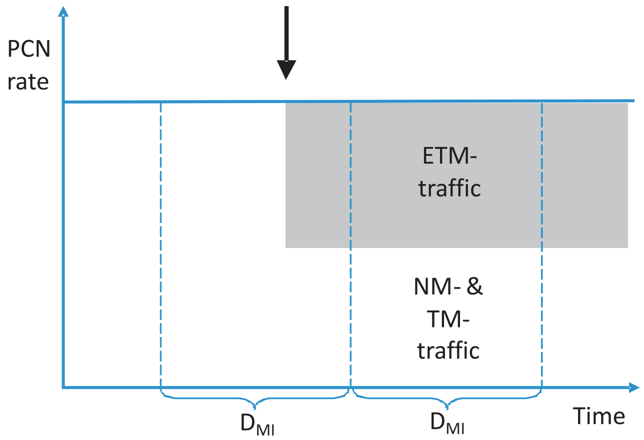
When the rate of ETM-traffic decreases like in Figures 3(b) and 3(d), the egress node is likely to overestimate EMR. As a result, the ingress node also overestimates the termination rate which holds for both dual and single marking. With dual marking, the termination rate is calculated by $TR = EMR$ and with single marking by $TR = \max(0, EMR - (u-1) \cdot NMR)$. As NMR does not decrease in the same way as EMR through the removal of SR-overload, MRT-DTR with single marking causes less overtermination than MRT-DTR with dual marking when the EMR is overestimated. When flows within the observed IEA are terminated, the ETM-traffic rate decreases like in Figure 3(b). This source of overtermination can be eliminated by enforcing a minimum inter-termination time (ITT) between two consecutive termination steps. The minimum ITT must cover at least the time to terminate a flow (flow termination time, FTT), one round trip time (RTT) from the ingress to the egress and back, and the duration of one measurement interval D_{MI}, i.e., $ITT = FTT + RTT + D_{MI}$. The latter is needed to avoid that termination uses an egress node's measurement report that still covers traffic from previously terminated flows. In Section 4.3 we show how larger ITTs avoid overtermination if the ETM-traffic rate decreases like in Figure 3(d) where traffic from other IEAs has been terminated. Other causes for the removal of SR-overload like rerouted traffic flapping back to its primary paths can also be sources for this type of overtermination, but they are difficult to eliminate.



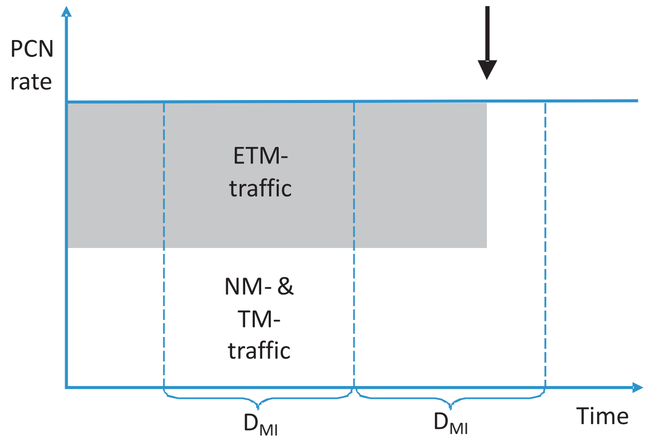
(a) Increase of ETM- and non-ETM-traffic rates for an IEA due to rate increase (measurement at egress).



(b) Decrease of ETM- and non-ETM-traffic rates for an IEA due to rate reduction (measurement at egress).



(c) Increase of ETM- and decrease of non-ETM-traffic rates for an IEA while the overall PCN traffic rate remains stable (measurement at egress).



(d) Decrease of ETM- and increase of non-ETM-traffic rates for an IEA while the overall PCN traffic rate remains stable (measurement at egress).

Figure 3: Measurement intervals with incipient and ceasing SR-pre-congestion lead to non-representative estimations of ETM- and non-ETM-traffic rates for IEAs.

4.2.2. Analysis of MRT-ITR

We consider MRT-ITR with dual and single marking when the rate of ETM-traffic increases like in Figures 3(a) and 3(c). When the ingress node receives a measurement report from the egress node, it first examines it for SR-pre-congestion. With dual marking, $EMR > 0$ is a sign for SR-pre-congestion while with single marking $u \cdot NMR < NMR + EMR$ indicates SR-pre-congestion. Note that single marking possibly cannot recognize incipient SR-pre-congestion if the measured EMR is too small which delays the termination process. If the ingress node recognizes SR-pre-congestion, it starts the measurement of the sent PCN ingress rate IR. When the measured IR is available, the ingress node calculates the termination rate by $TR = IR - SAR$ with the sustainable aggregate rate $SAR = NMR + TMR$. The ingress node is likely to under- or overestimate SAR based on the data of the first measurement report indicating SR-pre-congestion. Therefore, the ingress node should use the data from the second measurement report which provides a more accurate value for SAR. This report normally has arrived already at the end of the measurement interval of IR so that this rule does not induce additional delay for the termination process. Then, the ingress node terminates an appropriate set of flows to reduce the PCN traffic rate of the IEA by TR, but only if the new report still indicates SR-pre-congestion.

If the PCN traffic rate increases during the measurement of IR at the ingress node like in Figure 3(a), the IR is likely to be underestimated as well as TR so that the ingress node possibly terminates too little traffic and another termination step is needed.

When the rate of non-ETM-traffic decreases like in Figure 3(b) where the observed IEA has terminated traffic, then the ingress node overestimates $SAR = NMR + TMR$. This possibly – but not necessarily – leads to undertermination. In contrast, when the rate of non-ETM-traffic increases like in Figure 3(d) because other IEAs have reduced their traffic on the shared bottleneck link, then the ingress node possibly underestimates $SAR = NMR + TMR$. This is likely to cause overtermination because the sustainable aggregate rate SAR is lower than the ingress rate IR measured by the ingress node. If the rate reduction of the other IEAs is due to a termination event, sufficiently long ITTs can help to avoid overtermination (see Section 4.3). As mentioned above, rerouted traffic of other IEAs flapping back to the primary path can also reduce traffic on the bottleneck link, but this source of overtermination is rather difficult to eliminate.

When the sent overall PCN traffic rate decreases within a measurement interval at the ingress node like in Figure 3(b), the ingress node overestimates IR and TR and terminates too much traffic. If the rate decrease is due to a termination event of the considered IEA, overtermination can be avoided by starting the measurement interval only after all previous termination steps are finished. This leads to a minimum inter-termination time $ITT = FTT + D_{MI}$.

4.3. Impact of Multiple IEAs with Different RTTs

We consider multiple IEAs on a SR-pre-congested link and show that overtermination can occur when the IEAs have different RTTs. This phenomenon has been reported first in [40]. We quantify the strength of potential overtermination and propose a method to avoid it. We consider only MRT-ITR with dual marking in our analysis, but the results also apply to MRT-DTR and to single marking.

4.3.1. Experiment Setup

We consider the setting in Figure 4 with two ingress nodes A_0 and A_1, one interior node B, and one egress node C. The IEAs from A_0 and A_1 to C are called IEA_0 and IEA_1. IEA_0 is carried over B to C and IEA_1 is usually carried directly to C. However, due to a failure of the direct link from A_1 to C, IEA_1 is rerouted over B to C. RTT_0 is the RTT from A_0 over B to C and back, and RTT_1 is the RTT from A_1 over B to C and back. We assume in our example that RTT_0 is larger than RTT_1. When IEA_1 is rerouted, SR-overload possibly occurs on the link l between B and C. In the following we focus on this link. Its admissible rate is AR_l and its supportable rate is $SR_l = u \cdot AR_l$. The parameter u is actually needed for single marking only, but we use it also for dual marking to control the size of $SR_l = u \cdot AR_l$ in our experiments.

Figure 5 shows a time diagram for the termination process. When egress node C detects the SR-overload caused on link l by the arrival of ETM-packets, it starts continuously measuring the rates NMR_i, TMR_i, and EMR_i for $i \in \{0,1\}$, and sends these values at the end of the measurement intervals to A_0 and A_1, respectively. Ingress node A_i sees that EMR_i is larger than zero and measures the sent PCN ingress rate IR_i. At the end of the measurement interval, it calculates the termination rate by $TR_i = IR_i - SAR_i$ with $SAR_i = NMR_i + TMR_i$ using the latest values for NMR_i and TMR_i. Then, it terminates an appropriate number of flows. Since IEA_1 has a shorter RTT than IEA_0, the termination effect of A_1 is visible earlier than the one of A_0 both at link l and at egress node C. When the effect of A_1's termination is visible at the link l, the SR-overload is not yet fully removed until the effect of A_0's termination is visible, too. Within that time, some traffic of IEA_1 is still ETM although the rate of IEA_1 has already been sufficiently reduced. As a result, ingress node A_1 underestimates the sustainable aggregated rate SAR_1 of IEA_1 and performs another termination step which finally leads to overtermination.

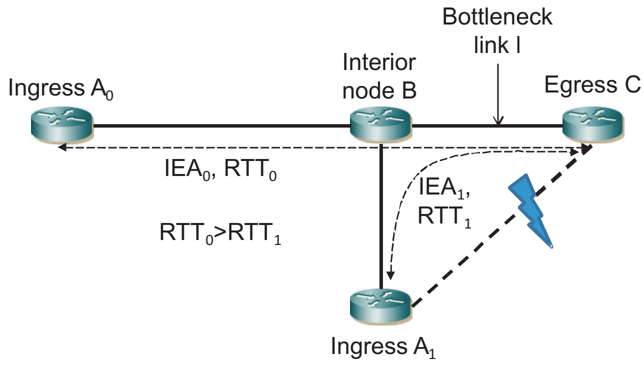


Figure 4: IEA_1 is rerouted and causes SR-overload on link l.

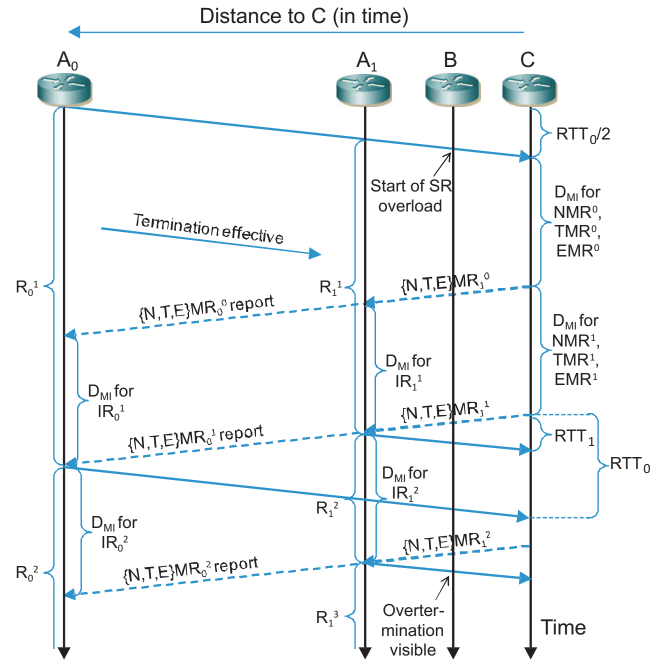


Figure 5: Time diagram: different RTTs for IEA_0 and IEA_1 lead to asynchronous termination and possibly to overtermination. The variables are explained in Section 4.3.2.

4.3.2. Analysis

We propose an analysis to quantify the presented kind of overtermination under challenging conditions. First, we explain the considered networking scenario and clarify some notation. The measurement intervals at the ingress and egress nodes are D_{MI} long. The measurement intervals at egress node C are numbered by $j = 0, 1, \ldots$, starting with the one that covers SR-overload for the first time. Corresponding measured rates are denoted NMR_i^j, TMR_i^j, and EMR_i^j for IEA_i. The measurement intervals at the ingress nodes are numbered by $m = 1, 2, \ldots$ and the measured sent PCN ingress rates are denoted IR_i^m. At the end of these measurement intervals, the ingress nodes possibly terminate traffic and the corresponding termination step is numbered by m. The rates of IEA_i before potential termination step m are named R_i^m. We assume in our setting $R_0^1 = AR_l = 1/u \cdot SR_l$, i.e., AR_l is fully utilized by the

PCN traffic of IEA_0. We choose the initial rate R_1^1 of IEA_1 so that it causes a relative SR-overload of q on the bottleneck link l after reroute. A value of $q = 0$ means no SR-overload. Hence, we have $R_1^1 = (1+q) \cdot SR_l - R_0^1 = (1+q-1/u) \cdot SR_l$. For the sake of simplicity, we assume that ingress nodes immediately terminate flows after heaving computed TR_i^m. That means, the flow termination time (FTT) is zero so that ingress nodes can start the measurement of IR_i^{m+1} immediately after the one of IR_i^m if needed.

We now analyze the termination process. We assume $0 \leq RTT_1 \leq RTT_0 \leq D_{MI}$ to simplify the analysis. Immediately after the reroute, the initial rates R_0^1 and R_1^1 cause SR-overload on the common bottleneck link l so that only the fraction $\frac{SR_l}{R_0^1+R_1^1} = \frac{1}{1+q}$ of the PCN traffic remains non-ETM. As soon as egress node C sees the first ETM-packet, it starts measurement interval $j = 0$. We choose this optimization of MRT-DTR (see Section 4.2.1) to simplify our analysis, otherwise the start of the measurement interval is random. The resulting measured rates are $NMR_i^0 = 0$, $TMR_i^0 = R_i^1 \cdot \frac{SR_l}{R_0^1+R_1^1}$, and $EMR_i^0 = R_i^1 - TMR_i^0$. The egress node C sends them to the ingress nodes A_0 and A_1, and continues measuring. The ingress nodes A_0 and A_1 receive the measurement reports and measure IR_i^1. In the meanwhile, the ingress nodes receive from egress node C another measurement report with NMR_i^1, TMR_i^1, and EMR_i^1, which resemble very much the previous ones since no traffic has been terminated, yet. The ingress nodes calculate the sustainable aggregate rate

$$SAR_i^1 = NMR_i^1 + TMR_i^1 = R_i^1 \cdot \frac{SR_l}{R_0^1 + R_1^1} \qquad (2)$$

and terminate $TR_i^1 = IR_i^1 - SAR_i^1$ traffic. The effect of both termination steps becomes visible at the egress node $RTT_i + 2 \cdot D_{MI}$ time after egress node C observed the first ETM-packet, i.e., in the third considered measurement interval which has number $j = 2$. The newly measured rates are reported to the ingress nodes A_0 and A_1. As $EMR_i^2 > 0$, the ingress nodes calculate SAR_i^2. If the RTTs of both IEAs are the same, then SAR_i^2 equals SAR_i^1 for $i \in \{0,1\}$ so that no additional termination step is performed provided that enough traffic has been removed in the first termination step. However, if we have $RTT_0 > RTT_1$, then $SAR_0^2 > SAR_0^1$ and $SAR_1^2 < SAR_1^1$ hold so that A_1 terminates traffic again. The exact value for SAR_1^2 can be calculated as follows:

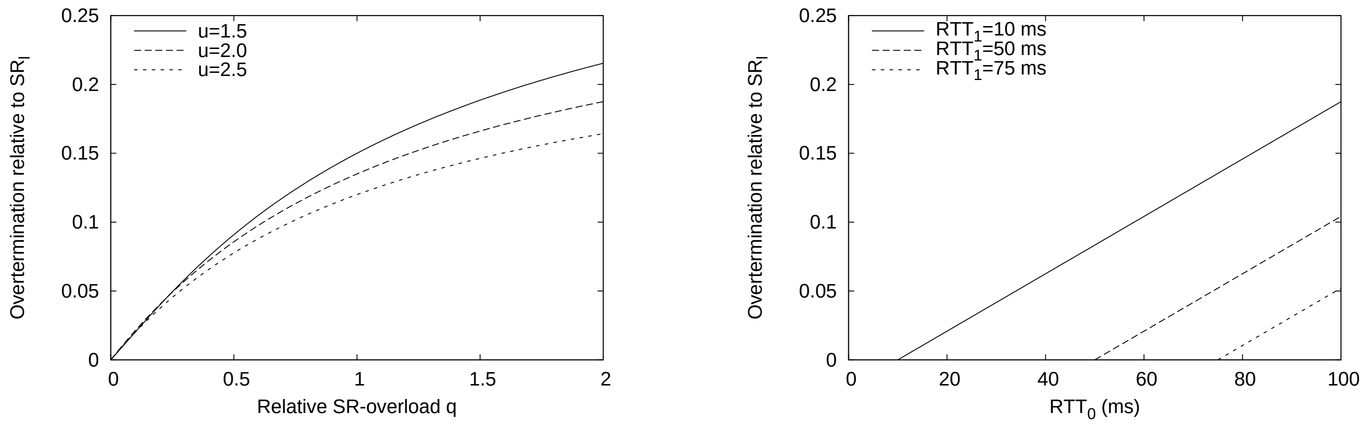
$$\begin{aligned} SAR_1^2 &= NMR_1^2 + TMR_1^2 \qquad (3)\\ &= \frac{1}{D_{MI}} \cdot \Bigg(RTT_1 \cdot R_1^1 \cdot \frac{SR_l}{R_0^1 + R_1^1} + \\ &\quad (RTT_0 - RTT_1) \cdot R_1^2 \cdot \frac{SR_l}{R_0^1 + R_1^2} + \\ &\quad (D_{MI} - RTT_0) \cdot R_1^2 \cdot \frac{SR_l}{R_0^2 + R_1^2} \Bigg) \end{aligned}$$

whereby the rates R_i^2 equal SAR_i^1. This equation basically weights the PCN traffic of IEA_1 observed by the egress node in the third measurement interval with the different probabilities for non-ETM-packets experienced on link l. After A_1's second termination step is visible at the bottleneck link l, the relative overtermination on that link is $OT = \frac{SAR_1^1 - SAR_1^2}{SR_l}$.

4.3.3. Analytical Results

We quantify the caused overtermination for measurement intervals of duration $D_{MI} = 100$ ms. Figure 6(a) shows it for $RTT_0 = 100$ ms, $RTT_1 = 10$ ms, and different values of u and relative SR-overload q. Overtermination strongly increases with the relative SR-overload q and is larger for smaller u-values that control the relation between AR_l and SR_l. Overtermination in the order of 15% – 20 % can be easily achieved in this setting.

Figure 6(b) illustrates the overtermination for a relative SR-overload $q = 2.0$, $SR_l = 2.0 \cdot AR_l$ ($u = 2.0$), and different RTT_0 and RTT_1. Overtermination increases about linearly with RTT_0 and is smaller for larger RTT_1. The overtermination effect vanishes if RTT_0 and RTT_1 are equally long.



(a) $RTT_0 = 100$ ms, $RTT_1 = 10$ ms.

(b) Relative SR-overload $q = 2.0$, $u = \frac{SR_l}{AR_l} = 2.0$.

Figure 6: Overtermination on link l relative to SR_l.

4.3.4. *Prevention of Overtermination due to Different RTTs*

We propose a method to avoid or reduce overtermination that is due to different RTTs. Overtermination can be prevented if subsequent termination steps are delayed until new PCN feedback reflects the effects of all previous terminations. We derive the appropriate inter-termination time for MRT-DTR and MRT-ITR.

The presented example and the analysis apply for MRT when a common egress node starts measurement intervals with the receipt of the first ETM-packet. However, the egress nodes for IEAs sharing a common bottleneck link may be different and they may measure PCN feedback periodically. As a consequence, ETM-packets may be reported to the ingress node already shortly after their arrival at the egress node or almost D_{MI} time later. The ITT needs to account for that uncertainty. The interval starting with the egress node sending the PCN feedback until the effect of the termination becomes visible at the egress node is $\max_i(RTT_i + FTT_i)$ for MRT-DTR and $\max_i(RTT_i + FTT_i) + D_{MI}$ for MRT-ITR. Then, the ongoing measurement interval at the egress node must be finished before PCN feedback may be collected. This adds another D_{MI} delay. Finally, the actual data collection takes another D_{MI} time. Hence, to avoid overtermination due to different RTTs, MRT-DTR requires $ITT = \max_i(RTT_i + FTT_i) + 3 \cdot D_{MI}$ and MRT-ITR requires $ITT = \max_i(RTT_i + FTT_i) + 4 \cdot D_{MI}$.

4.4. *Impact of Packet Loss and Packet Drop Policies*

Packet loss reduces the rates of NM-, TM-, or ETM-traffic received by the PCN egress node. ETM- or non-ETM-packets may be preferentially dropped, or packets may be dropped independently of their markings. We show that the packet drop policy affects the termination process of MRT-DTR and MRT-ITR in different ways.

4.4.1. *Experiment Setup*

We assume that packet loss inside a node occurs before packets are metered and marked. Therefore, ETM-packets can be lost only at a downstream node relative to the node which marked them with ETM. Hence, two SR-pre-congested links are needed to provoke a situation where ETM-packets are lost: one SR-pre-congested link that marks packets with ETM and another SR-pre-congested link that even drops PCN packets.

To keep things simple, we consider the experiment setup depicted in Figure 7. A single IEA with initial 25 Mbit/s is transmitted over the two adjacent links l_0 and l_1. The configured

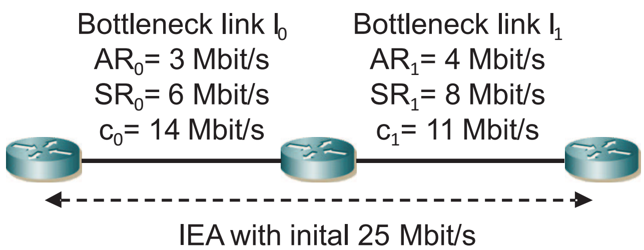


Figure 7: Experiment setup to evaluate the impact of packet loss.

admissible and supportable rates AR_i and SR_i of link l_i as well as its capacity c_i are given in the figure. We chose the values in the experiment so that all interesting phenomena can be shown with a single parameter set.

We study the reduction of the PCN rate of the IEA due to termination for MRT-DTR and MRT-ITR, dual and single marking, and for the three packet drop policies: drop ETM-packets (DEP), drop non-ETM packets (DNP), and drop random packets (DRP). We assume that DRP drops the same fraction of ETM- and non-ETM-traffic.

4.4.2. *Analysis*

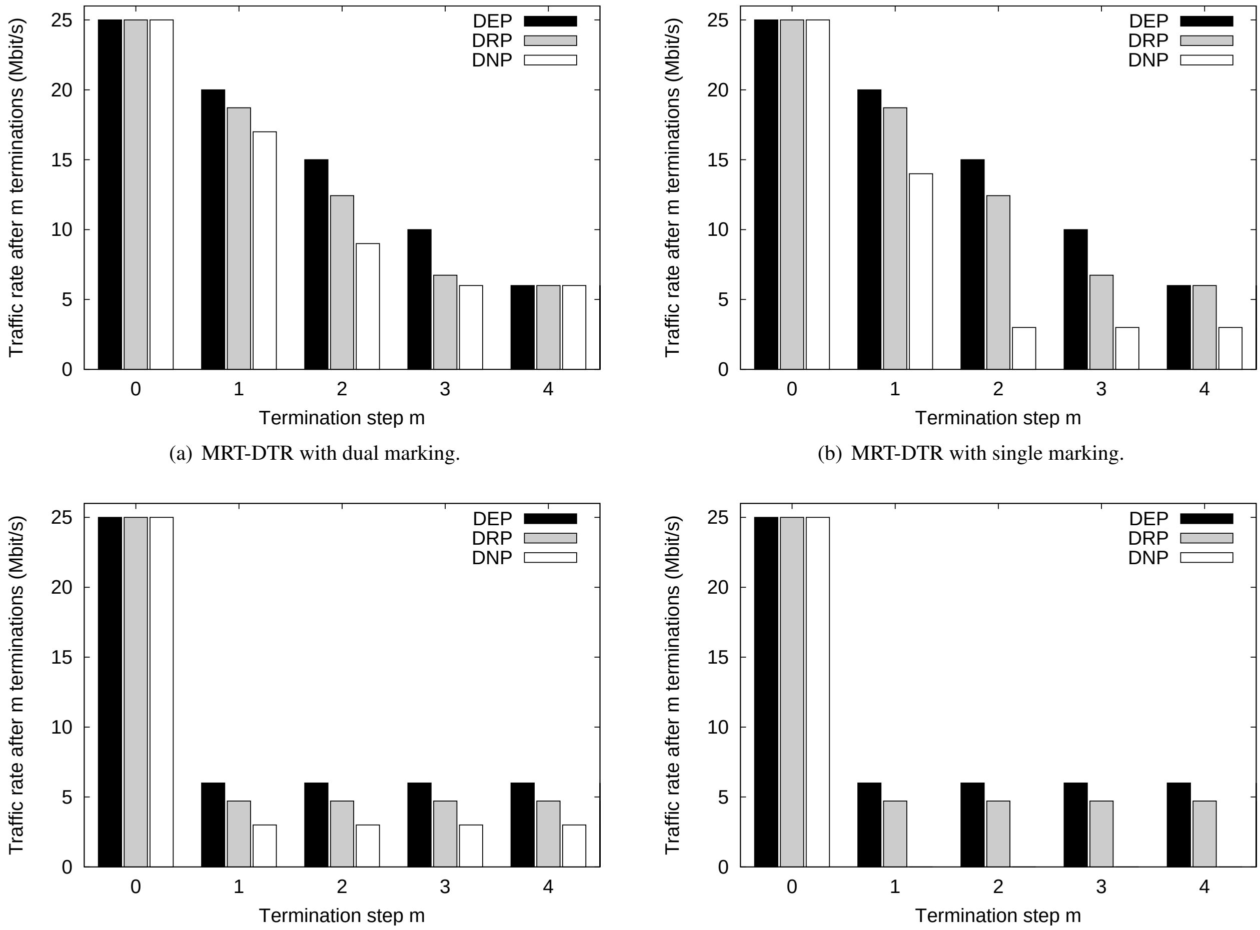
To investigate the termination process, we use a step-by-step analysis, i.e., we calculate the rates of ETM- and non-ETM-traffic of the considered IEA on link l_0 before marking, on link l_0 after marking, on link l_1 after packet loss but before marking, and on link l_1 after marking. Based on that information, the rate of the IEA after the next termination step is calculated and the analysis is repeated with the new initial rate. This analysis is straightforward but cumbersome so that we do not show any equations. When cross traffic appears on multiple pre-congested links, a more sophisticated analysis is needed. Then, overtermination can occur for all termination methods even without packet loss [35]. However, this phenomenon is orthogonal to the observations reported in this section. The results of the analysis are summarized in Figures 8(a)–8(d) and discussed in the following.

4.4.3. *MRT-DTR with Dual Marking*

Figure 8(a) shows for MRT-DTR with dual marking the rate of the IEA after m termination steps. We observe that several termination steps are needed to reduce the PCN rate down to the expected 6 Mbit/s. In the absence of packet loss, the rate of ETM-traffic exactly corresponds to SR-overload and equals the amount of traffic that needs to be terminated. When ETM-packets are lost, the termination rate is underestimated and undertermination occurs so that additional termination steps are required. Since DEP loses more ETM-packets than DRP and DNP, the corresponding termination process takes longer for DEP. The question arises whether DEP possibly loses so much traffic that termination does not work anymore. The gap between the SR and the bandwidth c of a link determines the minimum amount of ETM-traffic that leaves the link with DEP in case of packet loss. Based on this difference, a lower bound for the termination speed can be calculated. As long as $SR < c$ holds for all links of a PCN domain, traffic is still terminated. Hence, DEP and DRP cannot prevent termination for MRT-DTR with dual marking, but they delay the termination process if several steps are needed to remove SR-overload.

4.4.4. *MRT-DTR with Single Marking*

Figure 8(b) illustrates the termination process for MRT-DTR and single marking. With DEP and DRP, the termination process is the same as for dual marking. However, in case of DNP, overtermination occurs as only 3 Mbit/s instead of the expected 6 Mbit/s PCN traffic remain after the second termination step.



(a) MRT-DTR with dual marking.

(b) MRT-DTR with single marking.

(c) MRT-ITR with dual marking.

(d) MRT-ITR with single marking.

Figure 8: PCN traffic rate after several termination steps for MRT-DTR and MRT-ITR with dual and single marking.

This happens because MRT-DTR with single marking calculates the termination rate by $TR = NMR + EMR - u \cdot NMR$ and if the rate of non-ETM-traffic NMR is too low, TR is overestimated which possibly leads to overtermination. This cannot happen with DEP. Overtermination neither occurs with DRP because it drops the same fraction of ETM- and non-ETM-traffic which just reduces the termination rate accordingly. Hence, MRT-DTR with single marking should be deployed only with DEP or DRP.

4.4.5. MRT-ITR with Dual Marking

Figure 8(c) shows the termination process for MRT-ITR and dual marking. The termination is already completed after a single termination step. We observe that overtermination occurs with DRP and DNP as only 4.7 Mbit/s and 3 Mbit/s instead of 6 Mbit/s PCN traffic remain after termination. This happens because DRP and DNP drop non-ETM-packets which leads to an underestimation of the sustainable aggregate rate SAR with MRT-ITR. As a consequence, the termination rate $TR = IR - SAR$ is overestimated and too much traffic is terminated. With DEP, overtermination does not occur since non-ETM-packets are not lost so that a correct estimate for SAR is obtained. Hence, MRT-ITR with dual marking works correctly only with DEP.

4.4.6. MRT-ITR with Single Marking

Figure 8(d) visualizes the termination process for MRT-ITR and single marking. Again, overtermination occurs in case of DRP and DNP for the same reason as with dual marking. DNP even fully removes the PCN traffic so that the figure misses the corresponding bars. This can also be achieved for dual marking when different parameter settings are chosen in the experiment. Hence, also MRT-ITR with single marking should be deployed only with DEP.

4.5. Impact of Packet Loss on the Number of Required Termination Steps for MRT-DTR

In the absence of packet loss, MRT-DTR requires only a single termination step to remove SR-overload. However, in Section 4.4 we have shown that MRT-DTR needs multiple termination steps to fully remove SR-overload in the presence of packet loss. This delays the termination process and is the major disadvantage of MRT-DTR compared to MRT-ITR. We analytically calculate the number of required termination steps to remove

SR-overload and discuss the results. They are valid for MRT-DTR with dual and single marking.

4.5.1. Analysis

We consider a single link with bandwidth c and supportable rate SR. The link is faced with so much PCN traffic that a PCN packet loss probability of p occurs. The overall PCN traffic rate offered to the link can be written as $\frac{c}{1-p}$ and the overall rate to be terminated is then $\frac{c}{1-p} - SR$. In a single termination step, $c - SR$ traffic can be terminated. Therefore, the number of required termination steps m to fully remove SR-overload in the presence of an initial packet loss p is

$$m = \left\lceil \frac{\frac{c}{1-p} - SR}{c - SR} \right\rceil = \left\lceil \frac{\frac{1}{1-p} - \frac{SR}{c}}{1 - \frac{SR}{c}} \right\rceil. \quad (4)$$

Since packet loss is not an intuitive measure for SR-overload, we also consider the initial relative SR-overload q, i.e., the initial SR-overload in multiples of SR. Then, then number of required termination steps is

$$m = \left\lceil \frac{q \cdot SR}{c - SR} \right\rceil = \left\lceil \frac{q}{\frac{c}{SR} - 1} \right\rceil. \quad (5)$$

4.5.2. Analytical Results

Figure 9(a) shows the number of required termination steps for a relative supportable rate $\frac{SR}{c}$ and a given initial packet loss p. The diagram is partitioned by the lines into several areas that indicate the number of required termination steps m for $(\frac{SR}{c}, q)$ combinations belonging to that area. A single termination step suffices only in the absence of packet loss to fully remove SR-overload. Therefore, $m = 1$ is not in the figure. For a given relative supportable rate $\frac{SR}{c}$, the number of required termination steps increases with the initial packet loss p. Conversely, the overload induced by a certain packet loss p requires more termination steps when the supportable rate SR is closer to the link bandwidth c. Thus, to achieve fast termination even in the presence of high packet loss, the supportable rate SR should be chosen low enough compared to the link bandwidth c.

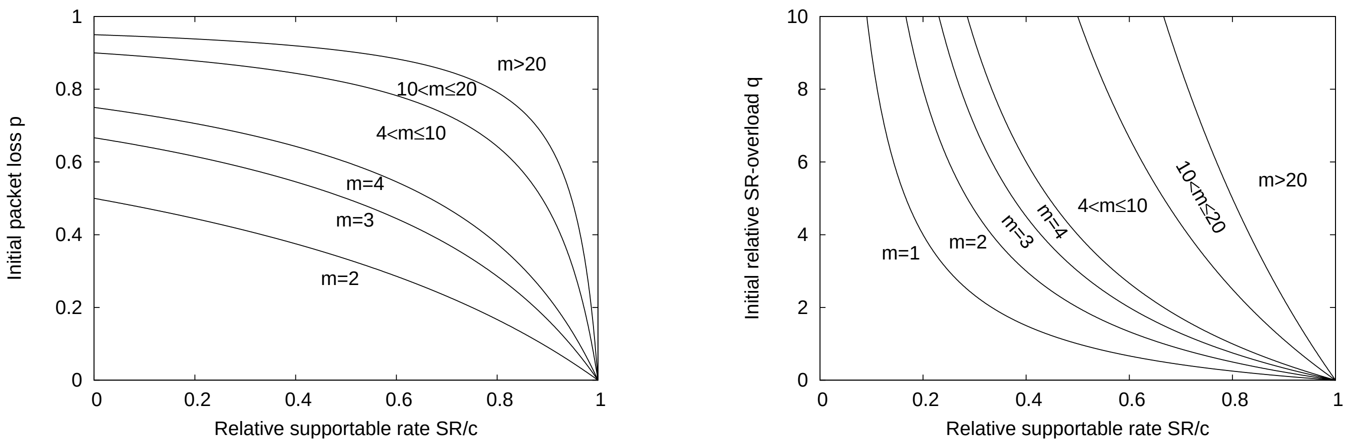
Figure 9(b) presents the same information in a different way. It indicates the number of required termination steps for combinations $(\frac{SR}{c}, q)$ of relative supportable rates $\frac{SR}{c}$ and the relative SR-overload q. A single termination step can remove an SR-overload that is significantly larger than SR if SR is small enough. For a relative supportable rate of $\frac{SR}{c} = 0.2$, SR-overload of up to 4 times SR can be terminated by two termination steps. In contrast, 4 termination steps are needed for a relative supportable rate of $\frac{SR}{c} = 0.8$ to remove a relative SR-overload of 100%. Hence, for MRT-DTR, there is a tradeoff between termination speed in the presence of high packet loss and the fraction of bandwidth that can be used to carry PCN traffic. The question whether MRT-DTR is fast enough boils down to the question whether surviving flows can afford a certain duration of QoS disruption, i.e. until SR-overload is removed, when many other flows are terminated.

4.6. Impact of a Small Number of Flows per IEA

PCN-based AC and FT are intended for networks with a sufficiently high PCN traffic rate per link [2]. This can be achieved when links carry a large number of small IEAs which is a likely scenario in future networks. If PCN domains are very large in terms of the number of ingress and egress nodes, only a very small number of realtime flows is expected per IEA [41]. Then, flow termination might have the following granularity problem. If MRT is expected to terminate 25% of the traffic of an IEA, but the IEA has only two flows, either 0 or 1 flow can be terminated. We propose several flow termination policies to handle this situation and investigate their impact using packet-based simulation.

4.6.1. Flow Termination Policies

We propose new flow termination policies.



(a) Initial load given as packet loss p.

(b) Initial load given as relative SR-overload q (SR-overload in multiples of SR).

Figure 9: The area in the figures indicates the number of required termination steps m for a scenario where the initial load and the relative supportable rate $\frac{SR}{c}$ are given.

- *Aggressive termination* terminates so many flows that their overall rate is at least the termination rate TR.
- *Careful termination* terminates a set of flows whose overall rate is at most TR.
- *Proportional termination* first terminates a set of flows whose overall rate is at most TR. Let the difference between TR and the rates of the terminated flows be ΔR. Then another flow f with rate r_f is chosen for potential termination as well as a random number $0 < y < 1$. If $y < \frac{\Delta R}{r_f}$ holds, flow f is terminated.
- *Safe termination* reduces the termination rate by some safety margin and then uses proportional termination to terminate that rate. The margin is given as a fraction $x \geq 0$ of the traffic that should remain after termination. For MRT-ITR this means that the ingress node calculates the termination rate by $TR = \max(0, IR - (1+x) \cdot SAR)$. With MRT-DTR and dual marking, the ingress node calculates the termination rate by $TR = \max(0, EMR - x \cdot (NMR + TMR))$. With MRT-DTR and single marking, the ingress node calculates the termination rate by $TR = \max(0, NMR + EMR - (1+x) \cdot u \cdot NMR)$.

4.6.2. Experiment Setup

We consider a single bottleneck link with a supportable rate of $SR = 12$ Mbit/s. Initially, it carries $n_{IEA} = 50$ IEAs and some time later $n_{IEA} = 100$ IEAs due to a rerouting event. Each of the IEAs has $n_{IEA}^{flows} = 2$ flows with $r_f = 80$ kbit/s at simulation start. Then, 16 Mbit/s run over the bottleneck link which corresponds to an SR-overload of 33%. Hence, 25% of the flows should be removed so that only 12 Mbit/s PCN traffic remain on the bottleneck link. However, each IEA can remove either 0, 1, or 2 flows. Thus, there is a granularity problem.

We use a packet-based simulation to study the time-dependent PCN traffic rate on the bottleneck link. We assume periodic voice traffic with constant packet inter-arrival times $IAT = 20$ ms and constant packet sizes $B = 200$ bytes. To avoid simulation artifacts due to overly exact arrival times, we add some uniformly distributed jitter to the packet transmission times of at most $D_{pkt}^{max} = 1$ ms. The excess marker on the bottleneck link is configured with reference rate SR and a bucket size of $0.05 \text{ s} \cdot SR$, i.e. 0.6 Mbit. The measurement intervals are $D_{MI} = 100$ ms long. We run 100 simulations and average the obtained time-dependent traffic rates. The 95% confidence intervals are smaller than 1% of the obtained mean values. We omit the confidence intervals in the figures for the sake of clarity.

We simulate MRT-DTR with dual marking where the egress node restarts the measurement of PCN feedback with the receipt of the first ETM-packet. We obtain almost the same results for MRT-ITR and dual marking when the egress node periodically measures PCN feedback; in that case, the termination process is at most D_{MI} time delayed. Similar results also apply for single marking. However, the results for single marking are overlaid by additional problems that are studied in Section 4.7.

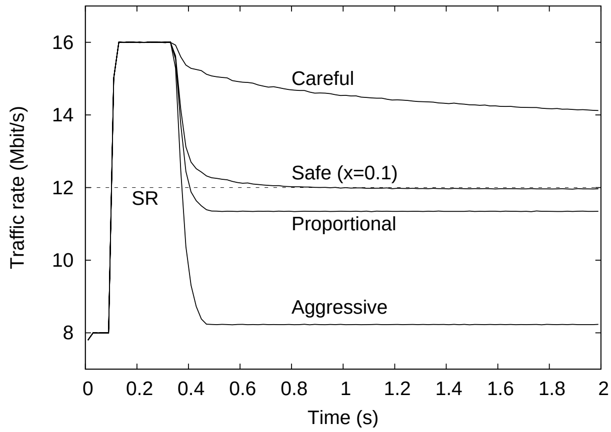


Figure 10: Time-dependent PCN traffic rate on the bottleneck link with MRT-DTR and dual marking: various flow termination policies may cause over- or undertermination.

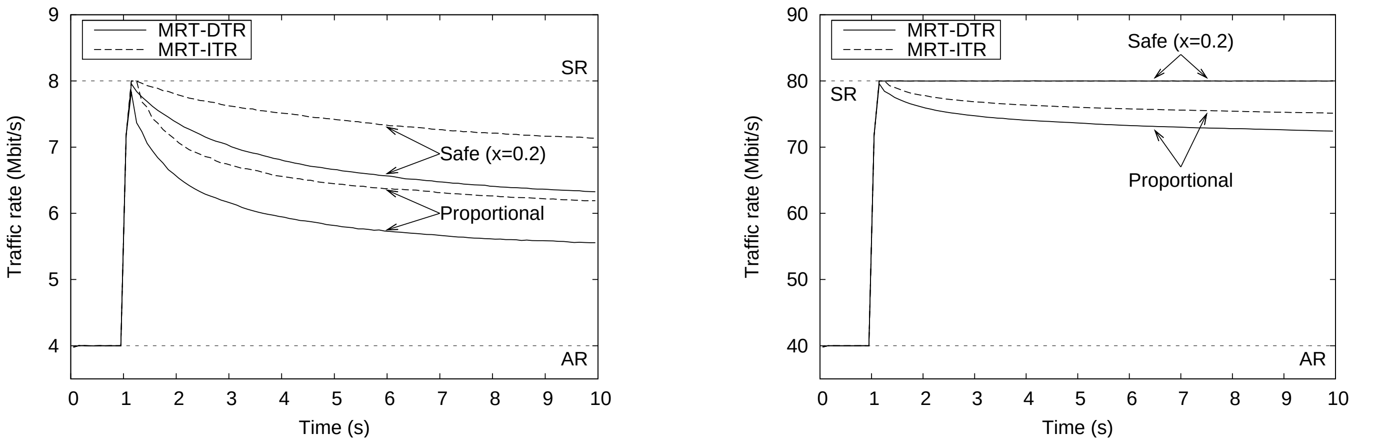
4.6.3. Simulation Results

Figure 10 shows how the PCN traffic rate on the SR-pre-congested link evolves with the four different flow termination policies. Aggressive termination leads to significant overtermination. After termination only 8 Mbit/s out of the 16 Mbit/s remain on the link because every IEA removes one flow which corresponds to 50% termination instead of the required 25% termination. This is an overtermination of 33%. Careful termination leads to significant undertermination on the bottleneck link because it does not terminate any flow on most IEAs. As the number of ETM-packets per IEA is subject to statistical fluctuations, the amount sometimes suffices that an IEA terminates a flow. Proportional termination mostly terminates no or one flow per IEA. The figure shows that the PCN traffic rate on the bottleneck link is reduced to a bit less than the desired *SR*. Safe termination with a margin of 10% terminates exactly as much traffic as needed so that the PCN traffic rate eventually meets the desired *SR* on the bottleneck link. Thus, proportional or safe termination should be used in practice to avoid over- and underterminaiton in the presence of a small number of flows per IEA.

Another aspect is fairness for which we do not provide any simulation data. Different IEAs may receive different rates of ETM-traffic as PCN feedback which can lead to different fractions of terminated flows among IEAs. This is unfair and not desirable but acceptable in exceptional situations where traffic is terminated.

4.7. Impact of a Small Number of Packets per Measurement Interval

The number of ETM-packets per IEA is subject to statistical fluctuations. As single marking marks packets with ETM already in the presence of AR-overload, it is possible that the fraction of ETM-packets in a measurement interval is so large that flows are terminated even in the absence of SR-pre-congestion. We quantify this effect, propose countermeasures, and show their effectiveness.



(a) 50 PCN packets per measurement interval ($n_{IEA}^{flows} = 10$ voice flows per IEA).

(b) 500 PCN packets per measurement interval ($n_{IEA}^{flows} = 100$ voice flows per IEA).

Figure 11: Time-dependent PCN traffic rate on the bottleneck link for MRT-DTR and MRT-ITR with single marking: flows are terminated in the absence of SR-pre-congestion ($n_{IEA} = 10$ IEAs on the link after rerouting).

4.7.1. Experiment Setup

We use a similar simulation setup as in the previous section. Due to single marking instead of dual marking, the excess marker is configured with the admissible rate instead of the supportable rate. The simulation starts with $\frac{n_{IEA}}{2}$ IEAs on the bottleneck link. The resulting PCN traffic rate corresponds to the admissible rate of the link. After 1 s, additional $\frac{n_{IEA}}{2}$ IEAs are carried over the link which may happen due to a rerouting event. The supportable rate of the link is configured so that it corresponds to the rate of these $n_{IEA} \cdot n_{IEA}^{flows}$ flows, i.e., no flow needs to be terminated.

We consider two different versions of MRT-DTR. In the one version, the egress node restarts measuring PCN feedback with the first received ETM-packet and the ingress node terminates traffic as soon as signalled PCN feedback indicates termination. In the second version, the egress node periodically measures PCN feedback, but the ingress node terminates traffic only if the previous PCN feedback also required termination. We have proposed this second version also in Section 4.2.1. In the absence of packet loss, it leads to the same termination process as MRT-ITR. In the following, we denote the first MRT version by MRT-DTR and the second MRT version by MRT-ITR.

4.7.2. Simulation Results

Figure 11(a) shows the PCN traffic rate on the link for $n_{IEA}^{flows} = 10$ voice flows per IEA and two different flow termination policies. Initially, $\frac{n_{IEA}}{2} = 5$ IEAs are carried over the link, but after 1 s additional $\frac{n_{IEA}}{2} = 5$ IEAs appear due to rerouting. Therefore, AR-overload occurs, packets are marked with ETM, and flows are terminated. Flow termination happens in spite of the absence of SR-overload because the number of observed ETM-packets per measurement interval fluctuates and if it is sufficiently large, the ingress node terminates traffic. With proportional termination we observe overtermination of up to 30% for MRT-DTR and of up to 23% for MRT-ITR. The difference is due to the fact that MRT-ITR requires two consecutive PCN feedback per IEA indicating SR-overload to terminate traffic while for MRT-DTR a single PCN feedback indicating SR-overload is enough. Safety margins are intuitive countermeasures. However, safe termination with a large margin of 20% reduces overtermination only to about 20% for MRT-DTR and to 10% for MRT-ITR which is still not acceptable. The experiment is designed such that a measurement interval initially covers 50 PCN packets. The severity of the problem diminishes with an increasing number of PCN packets per measurement interval. Figure 11(b) illustrates the termination process with 10 times more flows per IEA, i.e. with 500 PCN packets per measurement interval. Proportional termination still leads to about 10% overtermination for MRT-DTR and to 6% for MRT-ITR, but safe termination with 20% safety margin fully avoids it.

Nearly the same relative evolution of the PCN traffic rate can be observed with $n_{IEA}^{flows} = 2$ and $n_{IEA}^{flows} = 20$ video flows (without figures). The packet inter-arrival time of these flows is 4 ms so that the experiment with video traffic leads to 50 and 500 PCN packets per measurement interval like in the experiment with voice traffic. As the packet size is set to 1500 instead of 200 bytes, the overall rate on the considered link carries 60 and 600 Mbit/s instead of 8 and 80 Mbit/s and c_l, AR_l, and SR_l are adapted accordingly in the simulation runs. The fact that almost the same relative evolution of the time-dependent PCN traffic rates is obtained shows that the observed overtermination is due to a low number of packets per measurement interval and not due to a low number of flows or a small traffic rate per IEA. Thus, another method to reduce potential overtermination is the prolongation of the measurement interval. This increases the number of PCN packets per measurement interval, but it also leads to a larger termination delay which is again undesirable.

These overtermination phenomena can be observed in simulations only if multiple IEAs are concurrently carried over a link. When only a single IEA is simulated, the ratio of the measured *NMR* and *EMR*, which are reported to the ingress node, is stable, AR-pre-congestion is correctly recognized, and flows are not unintentionally terminated. With multiple IEAs carried over a bottleneck link, PCN packets are marked with ETM on the pre-congested link independently of whether preceding packets of the same IEA have recently been marked with ETM. This leads to fluctuations of NMR and EMR which are a prerequisite for the observed overtermination. Furthermore, care must be taken to avoid that overly periodic packet transmissions lead to combinatoric effects and simulation artifacts. With dual marking, the reported problem cannot occur because packets become ETM only in the presence of SR-overload. Hence, termination cannot be triggered in the absence of SR-overload.

4.8. Impact of Multipath Routing

Multipath routing is frequently applied in IP networks in the form of the equal-cost multipath (ECMP) option [42]. Therefore, the applicability of PCN to networks with multipath routing is an important issue. The termination decisions of MRT methods are based on rate measurements of differently marked PCN traffic per IEA. This information is used to infer the pre-congestion state of the path belonging to the IEA which is meaningful only in case of single-path routing. In case of multipath routing, the obtained feedback stems from all partial (parallel) paths of the multipath carrying active flows. In addition, there is no information about which flows of an IEA are carried over an SR-pre-congested partial path and are candidates for termination. As a result, MRT with dual marking causes overtermination in case of multiple partial paths. MRT with single marking may cause not only overtermination but also undertermination, i.e., SR-pre-congestion is possibly not detected or not fully removed. In the following, we derive a mathematical model to quantify these effects of over- and undertermination and illustrate them for MRT with dual and single marking. The analysis and its results are valid for both MRT-DTR and MRT-ITR.

4.8.1. Analysis

We model the termination process assuming equal flow rates and denote the admitted traffic by the number of flows. The model state $s = (s_0, ..., s_{k-1})$ $(0 \leq i < k)$ indicates the number of current flows on k partial paths of an IEA. Admissible or supportable rates are assigned to links within a PCN domain, but in our analysis AR_i and SR_i indicate the number of admissible and supportable flows on each partial path. In reality, several flows are removed simultaneously at the end of each measurement interval. Our model neglects the time component which is here not of interest. Flows of an IEA are successively randomly chosen for termination and removed. The probability that a flow from path i is chosen for termination is $p(s,i) = \frac{s_i}{\sum_{0 \leq j<k} s_j}$ which yields the probability for the transition steps of a simple stochastic process

$$(s_0, ..., s_i, ..., s_{k-1}) \xrightarrow{p(s,i)} (s_0, ..., s_i - 1, ..., s_{k-1}) \qquad (6)$$

The process starts with $s_i = n_i$ flows. We compute the probability $p(s)$ of all states s with $0 \leq s_i \leq n_i$ by an iterative algorithm. The stop condition of the termination process depends on dual or single marking. In case of dual marking, the termination process stops if the SR-overload has been removed on all partial paths, i.e., if the condition

$$s_i \leq SR_i \quad \forall i : 0 \leq i < k \qquad (7)$$

is met. In case of single marking, the termination process stops if the overall received traffic rate is at most the rate of the non-ETM traffic ($\min(s_i, AR_i)$) multiplied by u, i.e., if the condition

$$\sum_{0 \leq i<k} s_i \leq u \cdot \sum_{0 \leq i<k} \min(s_i, AR_i) \qquad (8)$$

is met. The set $\mathcal{T}$ contains all states s in which the stochastic process terminates because the stop condition is met. The probability of the states in the terminating set $\mathcal{T}$ sums up to 1. Hence, we can calculate the average relative amount of overtermination and undertermination by

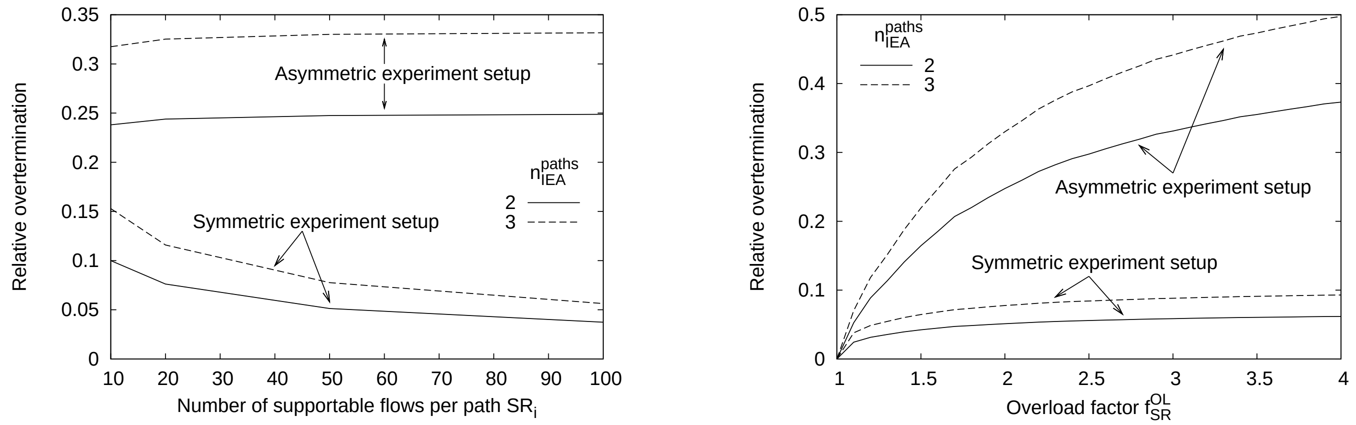
$$OT = \frac{\sum_{s \in \mathcal{T}} \sum_{0 \leq i<k} \max(0, \min(n_i, SR_i) - s_i) \cdot p(s)}{\sum_{0 \leq i<k} \min(n_i, SR_i)} \qquad (9)$$

$$UT = \frac{\sum_{s \in \mathcal{T}} \sum_{0 \leq i<k} \max(0, s_i - SR_i) \cdot p(s)}{\sum_{0 \leq i<k} \min(n_i, SR_i)}. \qquad (10)$$

4.8.2. MRT with Dual Marking and Multipath Routing

In this section we study MRT with dual marking and multipath routing. In case of SR-pre-congestion, flows are terminated from the IEA until no more ETM-packets arrive, i.e., until SR-overload is removed from all partial paths. Thereby, flows from non-SR-pre-congested partial paths are possibly also terminated and, therefore, overtermination occurs. We study the impact of several factors on overtermination and discuss signalling of additional information to reduce overtermination.

Impact of the Number of Flows per Partial Path. We perform the following symmetric experiment setup. An IEA carries traffic over $n_{IEA}^{paths} \in \{2, 3\}$ partial paths and each of them has the same supportable rate SR_i (in terms of number of flows) which is a variable parameter in our study. The initial number of flows $n_i = f_{SR}^{OL} \cdot SR_i$ is also the same on all partial paths and controlled by the overload factor $f_{SR}^{OL} = 2.0$. Figure 12(a) shows the analytically computed average overtermination after the termination process stopped depending on the number of supportable flows SR_i per partial path. With $n_{IEA}^{paths} = 2$ parallel paths per IEA, the average overtermination ranges between 4% and 10% and diminishes significantly with increasing numbers of supportable flows SR_i. With $n_{IEA}^{paths} = 3$ parallel paths per IEA, the average overtermination is larger but also decreases with increasing number of supportable flows SR_i. The figure also shows the results for an asymmetric experiment setup where only one partial path experiences an overload factor of $f_{SR}^{OL} = 2.0$ and the others are loaded with SR_i flows. In that case, the overtermination is about 25% for $n_{IEA}^{paths} = 2$ parallel paths per IEA and about 33% for $n_{IEA}^{paths} = 3$ parallel paths. In particular, the overtermination does not decrease with increasing numbers of supportable flows SR_i.



(a) Overload factor $f_{SR}^{OL} = 2.0$.

(b) Number of supportable flows per path $SR_i = 50$ flows.

Figure 12: Overtermination due to multipath routing with symmectric and asymmetric experiment setup.

Impact of the Overload Factor. We keep the number of supportable flows per partial path fixed at $SR_i = 50$ and vary the overload factor f_{SR}^{OL}. Figure 12(b) shows that the overtermination for the symmetric experiment is rather independent of the overload factor, lower than 6% for $n_{IEA}^{paths} = 2$ partial paths per IEA and lower than 10% for $n_{IEA}^{paths} = 3$. Thus, it has only minor impact. In contrast, in the experiment with only one SR-pre-congested partial path, the overtermination increases significantly with the overload factor f_{SR}^{OL} and reaches large values of up to 50%.

Impact of the Relative Size of the SR-Pre-Congested Path. We set the number of supportable flows per partial path on the non-SR-pre-congested paths to $SR_i = 50$. The overload factor for the SR-pre-congested path is $f_{SR}^{OL} = 2$ and we study the impact of the number of supportable flows on this path. The results are presented in Figure 13. The x-axis shows the supportable number of flows on the SR-pre-congested path relative to the other paths. For a relative size of $x = 1$ all partial paths have the same supportable rate and the observed overtermination equals the values in Figures 12(a) and 12(b) which are about 25% overtermination for $n_{IEA}^{paths} = 2$ partial paths per IEA and about 33% for $n_{IEA}^{paths} = 3$. When the SR-pre-congested partial path is smaller than the others, the overtermination can be significantly larger, i.e., 39% and 44% for $n_{IEA}^{paths} = 2$ and $n_{IEA}^{paths} = 3$ when the *SR* of the SR-pre-congested path is only 20% of the *SR* of its parallel paths. When the *SR* of the SR-pre-congested partial path is larger than the *SR* of its parallel paths, the overtermination can be significantly smaller.

Mitigating Overtermination by Additional Signalling. Overtermination due to multipath routing can be avoided for dual marking if egress nodes send information about flows with ETM-packets to the ingress nodes. As these flows are carried over SR-pre-congested paths, they are appropriate candidates for termination. If only a single partial path is SR-pre-congested, this method is obviously correct. If several partial paths are SR-pre-congested, this method helps to terminate traffic only from SR-pre-congested paths. However, overtermination can still occur in this case.

4.8.3. MRT with Single Marking and Multipath Routing

We illustrate over- and undertermination for MRT with single marking and multipath routing by analytical results and discuss signalling of additional information to improve the performance.

Analytical Results. In case of single marking, flows are terminated from the IEA until the fraction of ETM-packets is sufficiently small (see Equation (8)). This does not necessarily mean that SR-overload is removed from all partial paths. Thus, undertermination may occur. Note that one path may reveal overtermination and another undertermination after termination stops. Moreover, flows may not be terminated at all in spite of SR-pre-congestion on at least one partial path of the multipath since the IEA does not indicate SR-pre-congestion as Equation (8) is met. We perform some experiments that show how different but also how large the amount of over- and undertermination can be. We consider a single IEA with $n_{IEA}^{paths} = 2$ parallel paths, each of them having an admissible rate of $AR_i = 20$ flows, and $u = 2$. Thus, each partial path can carry up to 40 flows without being SR-pre-congested. We set the initial number of flows on the first partial path to $n_0 \in \{20, 40, 60\}$. Figure 14 shows the average relative over- and undertermination as well as their sum depending on the initial number of flows n_1 on the second partial path.

For $n_0 = 20$ initial flows on the first partial path, flows are not terminated for $n_1 \leq 60$ initial flows on the second partial path although the second partial path is already SR-pre-congested for more than $40 < n_1$ initial flows. Therefore, we observe up to 33% undertermination. For $n_1 > 60$, flows are terminated on both partial paths. With increasing n_1, undertermination decreases and overtermination increases, they occur simultaneously on both paths and sum up to about 33%. For $n_0 = 40$, none of the partial paths is SR-pre-congested for

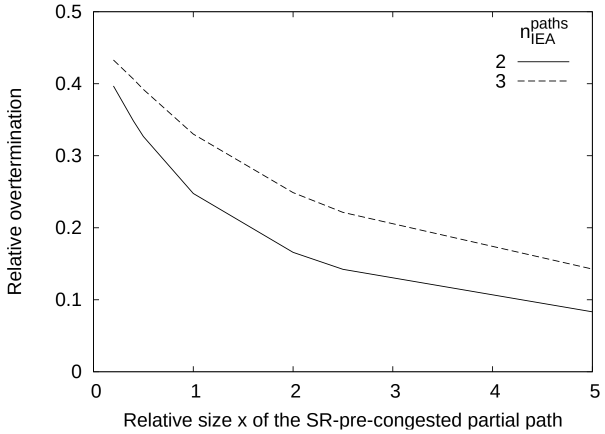


Figure 13: Overtermination due to multipath routing depending on the supportable rate of the SR-pre-congested partial path relative to the rate of the other non-SR-pre-congested partial paths ($f_{SR}^{OL} = 2.0$, $SR_i = 50$ flows).

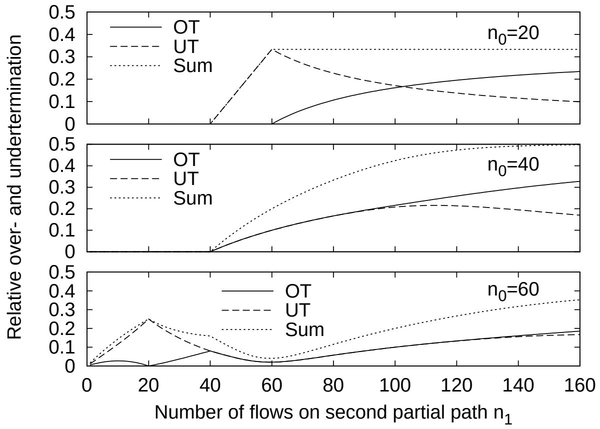


Figure 14: Average relative overtermination, undertermination, and their sum for single marking; $n_{IEA}^{paths} = 2$ parallel paths, $AR_0 = 20$, $AR_1 = 20$, $u = 2$, $n_0 \in \{20, 40, 60\}$, and n_1 varies.

$n_1 \leq 40$ and flows are not terminated. From $n_1 > 40$ on, SR-pre-congestion is indicated for the IEA and flows are terminated. The amount of over- and undertermination is the same up to a certain value of n_1. For $n_0 = 60$, the IEA indicates SR-pre-congestion for $n_1 < 20$ and $n_1 > 20$ and hence flows are terminated in these ranges. For $n_1 = 20$, the IEA does not indicate SR-pre-congestion although the first partial path is SR-pre-congested. For small values of $n_1 < 40$, there is more under- than overtermination. For $n_1 \geq 40$, the amount of over- and undertermination is the same up to a certain value when overtermination prevails. This is of course not an in-depth analysis, but the experiments show that over- and undertermination can be quite large and they are very sensitive to the load on the partial paths of a multipath.

Mitigating Overtermination by Additional Signalling. When only a single partial path is AR- or SR-pre-congested, overtermination can also be avoided with single marking. To that end, the egress node informs the ingress node about flows with ETM-packets. However, SR-overload is not necessarily detected so that undertermination may still occur. Furthermore, this method does not work when multiple partial paths are AR- or SR-pre-congested. ETM-packets can result from other AR-pre-congested paths whose flows should not be terminated. Therefore, it is not possible to reliably remove overtermination for single marking by additional signalling.

5. Summary

We have investigated three different flow termination methods that rely on measured PCN feedback: flow termination with directly measured termination rates (MRT-DTR), flow termination with indirectly measured termination rates (MRT-ITR), and flow termination with sustainable aggregate rates (MRT-SAR). They can be applied with dual and single marking.

In Section 4.1 MRT-SAR revealed to be extremely prone to overtermination when traffic descriptors are overestimated so that we excluded this method from further study. MRT-DTR and MRT-ITR suffer only from delayed termination when traffic descriptors are overestimated because then multiple termination steps are required.

In Section 4.2 we showed that incipient and ceasing SR-overload can lead to over- and underestimation of the rates of differently marked PCN traffic and to over- and undertermination. However, undertermination can be repaired by additional termination steps and overtermination can mostly be avoided by respecting sufficiently long inter-termination times and by calculating termination rates based on appropriate measurement reports. In Section 4.3 showed that overtermination can occur in particular if IEAs carried over a SR-pre-congested bottleneck link have significantly different RTTs. Sufficiently long ITTs again help to avoid overtermination.

In Section 4.4 we showed that packet loss can lead to overtermination. MRT-DTR with dual marking does not suffer from overtermination at all and is fastest when non-ETM-packets are preferentially dropped. MRT-DTR with single marking avoids overtermination when ETM-packets are preferentially dropped or when packets are dropped independently of their marking. MRT-ITR methods require preferential dropping of ETM-packets to avoid overtermination. While MRT-ITR can basically remove SR-overload in one shot, MRT-DTR requires several termination steps. Section 4.5 derived the number of required termination steps depending on various parameters.

Section 4.6 illustrated that extensive overtermination possibly occurs in the presence of IEAs with only a few flows because termination rates can be smaller than entire flows. We proposed the new proportional flow termination policy, possibly with a safety margin, that avoids this problem for dual marking. Section 4.7 shows that with single marking, traffic is already terminated in the presence of AR-pre-congestion without any SR-overload. The effect is significant when measurement intervals cover only a small number of PCN packets (~ 50). Proportional flow termination with a safety margin clearly reduces the overtermination but can hardly avoid it. Therefore, single marking is applicable only for IEAs with high traffic aggregation in terms of packets per second.

We demonstrated in Section 4.8 that all MRT methods – MRT-DTR or MRT-ITR with either dual or single marking – do not work well with multipath routing because the terminating ingress node does not know which flow of an IEA belongs to a SR-pre-congested path. Therefore, dual marking may lead to overtermination which can be mitigated when egress nodes signal information about marked flows to ingress nodes. Single marking may lead to both overtermination and undertermination in case of multipath routing and it cannot reliably detect and remove SR-overload under certain circumstances. This cannot be prevented by additional signalling.

MRT-ITR with preferential dropping of ETM-packets was adopted for standardization mainly because it terminates traffic faster than MRT-DTR. FT with single marking is simpler than FT with dual marking, but it possibly terminates flows without SR-pre-congestion and cannot be applied for networks with multipath routing. Therefore, both dual and single marking are currently standardized in IETF. MRT-ITR with dual marking is defined in [7] while MRT-ITR with single marking is standardized in [8].

6. Conclusion

Admission control (AC) and flow termination (FT) serve to achieve QoS for high priority traffic in the future Internet. Pre-congestion notification (PCN) is a load-dependent packet marking mechanism that supports simple feedback-based AC and FT for DiffServ domains. In this paper we have investigated multiple FT methods that are based on measured rates of differently marked PCN traffic. We documented pitfalls and challenging conditions that lead to overtermination and termination delay, thereby limiting the applicability of these methods. This leads to a better understanding of the tradeoffs in the design options and of PCN technology in general. We also proposed improvements to the FT algorithms to reduce overtermination under challenging conditions.

The current standardization process suggests FT with dual and single marking. Single marking is simpler from a technical and standardization point of view. However, FT with single marking causes overtermination in more situations than FT with dual marking. The results of this paper help operators to decide whether the simple FT with single marking satisfies their needs or whether they require the more complex FT with dual marking for their purposes.

Acknowledgements

The authors would like to thank Anna Charny, Bob Briscoe, Daisuke Satoh, Philip Eardley, and Joy (Xinyang) Zhang for fruitful discussions and valuable feedback.

Appendix

See Table 1.

Table 1: List of frequently used acronyms.

Acronym	Meaning
AC	admission control
AR	admissible rate
DEP	preferential dropping of ETM-packets
DNP	preferential dropping of non-ETM-packets
DRP	dropping of random packets
ECMP	equal-cost multipath
EMR	rate of ETM-traffic measured by the egress node
ETM	excess-traffic marked (PCN codepoint)
FT	flow termination
FTT	flow termination time
IEA	ingress-egress aggregate
IR	rate of PCN traffic sent and measured by the ingress node
ITT	inter-termination time
MRT	measured rate termination
MRT-DTR	MRT with directly measured termination rates
MRT-ITR	MRT with indirectly measured termination rates
MRT-SAR	MRT with measured sustainable aggregate rates
NM	not-marked (PCN codepoint)
NMR	rate of NM-traffic measured by the egress node
PCN	pre-congestion notification
RTT	round trip time
SAR	sustainable aggregate rate
SR	supportable rate
TM	threshold-marked (PCN codepoint)
TMR	rate of TM-traffic measured by the egress node
TR	termination rate

References

[1] S. Blake, D. L. Black, M. A. Carlson, E. Davies, Z. Wang, W. Weiss, RFC2475: An Architecture for Differentiated Services (Dec. 1998).
[2] P. Eardley (Ed.), RFC5559: Pre-Congestion Notification (PCN) Architecture (Jun. 2009).
[3] J. Wroclawski, RFC2211: Specification of the Controlled-Load Network Element Service (Sep. 1997).
[4] M. Menth, F. Lehrieder, B. Briscoe, P. Eardley, T. Moncaster, J. Babiarz, A. Charny, X. J. Zhang, T. Taylor, K.-H. Chan, D. Satoh, R. Geib, G. Karagiannis, A Survey of PCN-Based Admission Control and Flow Termination, accepted for IEEE Communications Surveys & Tutorials.
[5] Y. Bernet, P. Ford, R. Yavatkar, F. Baker, L. Zhang, M. Speer, R. Braden, B. Davie, J. Wroclawski, E. Felstaine, RFC2998: A Framework for Integrated Services Operation over Diffserv Networks (Nov. 2000).
[6] P. Eardley (Ed.), RFC5670: Metering and Marking Behaviour of PCN Nodes (Nov. 2009).
[7] A. Charny, F. Huang, G. Karagiannis, M. Menth, T. Taylor, PCN Boundary Node Behavior for the Controlled Load (CL) Mode of Operation, http://tools.ietf.org/html/draft-ietf-pcn-cl-edge-behaviour (Mar. 2010).
[8] A. Charny, J. Zhang, G. Karagiannis, M. Menth, T. Taylor, PCN Boundary Node Behavior for the Single-Marking (SM) Mode of Operation, http://tools.ietf.org/html/draft-ietf-pcn-sm-edge-behaviour (Mar. 2010).
[9] M. Menth, F. Lehrieder, Applicability of PCN-Based Admission Control, currently under submission.
[10] M. Menth, F. Lehrieder, PCN-Based Marked Flow Termination, currently under submission.
[11] S. Floyd, V. Jacobson, Random Early Detection Gateways for Congestion Avoidance, IEEE/ACM Transactions on Networking 1 (4) (1993) 397–413.
[12] B. Braden et al., RFC2309: Recommendations on Queue Management and Congestion Avoidance in the Internet (Apr. 1998).
[13] K. Ramakrishnan, S. Floyd, D. Black, RFC3168: The Addition of Explicit Congestion Notification (ECN) to IP (Sep. 2001).
[14] N. Spring, D. Wetherall, D. Ely, RFC3540: Robust Explicit Congestion Notification (ECN) (Jun. 2003).
[15] K. Nichols, S. Blake, F. Baker, D. L. Black, RFC2474: Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers (Dec. 1998).

[16] S. Floyd, RFC4774: Specifying Alternate Semantics for the Explicit Congestion Notification (ECN) Field (Feb. 2007).
[17] T. Moncaster, B. Briscoe, M. Menth, RFC5696: Baseline Encoding and Transport of Pre-Congestion Information (Nov. 2009).
[18] T. Moncaster, B. Briscoe, M. Menth, A PCN Encoding Using 2 DSCPs to Provide 3 Or More States, http://tools.ietf.org/html/draft-ietf-pcn-3-state-encoding (Feb. 2010).
[19] B. Briscoe, T. Moncaster, PCN 3-State Encoding Extension in a single DSCP, http://www.ietf.org/internet-drafts/draft-ietf-pcn-3-in-1-encoding-02.txt (Mar. 2010).
[20] M. Menth, J. Babiarz, T. Moncaster, B. Briscoe, PCN Encoding for Packet-Specific Dual Marking (PSDM), http://tools.ietf.org/html/draft-ietf-pcn-psdm-encoding (Jun. 2009).
[21] W. Almesberger, T. Ferrari, J.-Y. Le Boudec, SRP: A Scalable Resource Reservation for the Internet, Computer Communications 21 (14) (1998) 1200–1211.
[22] I. Stoica, H. Zhang, Providing Guaranteed Services without per Flow Management, in: ACM SIGCOMM, Boston, MA, 1999.
[23] R. Szábó, T. Henk, V. Rexhepi, G. Karagiannis, Resource Management in Differentiated Services (RMD) IP Networks, in: International Conference on Emerging Telecommunications Technologies and Applications (ICETA 2001), Kosice, Slovak Republic, 2001.
[24] R. J. Gibbens, F. P. Kelly, Resource Pricing and the Evolution of Congestion Control, Automatica 35 (12) (1999) 1969–1985.
[25] R. J. Gibbens, F. P. Kelly, Distributed Connection Acceptance Control for a Connectionless Network, in: 16^{th} International Teletraffic Congress (ITC), Edinburgh, UK, 1999, pp. 941 – 952.
[26] F. Kelly, P. Key, S. Zachary, Distributed Admission Control, IEEE Journal on Selected Areas in Communications 18 (12) (2000) 2617–2628.
[27] M. Karsten, J. Schmitt, Admission Control based on Packet Marking and Feedback Signalling – Mechanisms, Implementation and Experiments, Technical Report 03/2002, Darmstadt University of Technology (2002).
[28] M. Karsten, J. Schmitt, Packet Marking for Integrated Load Control, in: IFIP/IEEE Symposium on Integrated Management (IM), 2005.
[29] M. Menth, R. Martin, J. Charzinski, Capacity Overprovisioning for Networks with Resilience Requirements, in: ACM SIGCOMM, Pisa, Italy, 2006.
[30] S. Iyer, S. Bhattacharyya, N. Taft, C. Diot, An Approach to Alleviate Link Overload as Observed on an IP Backbone, in: IEEE Infocom, San Francisco, CA, 2003.
[31] M. Menth, F. Lehrieder, Comparison of Marking Algorithms for PCN-Based Admission Control, in: 14^{th} GI/ITG Conference on Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB), Dortmund, Germany, 2008, pp. 77–91.
[32] M. Menth, F. Lehrieder, Performance Evaluation of PCN-Based Admission Control, in: International Workshop on Quality of Service (IWQoS), Enschede, The Netherlands, 2008.
[33] S. Latre, B. Vleeschauwer, W. Van de Meerssche, S. Perrault, F. De Turck, P. Demeester, K. De Schepper, C. Hublet, W. Rogiest, S. Custers, W. Van Leekwijck, An Autonomic PCN based Admission Control Mechanism for Video Services in Access Networks, in: 4^{th} IEEE Workshop on Autonomic Communications and Network Management (ACNM, Long Island, NY, USA, 2009.
[34] F. Lehrieder, M. Menth, Marking Conversion for Pre-Congestion Notification, in: IEEE International Conference on Communications (ICC), Dresden, Germany, 2009.
[35] F. Lehrieder, M. Menth, PCN-Based Flow Termination with Multiple Bottleneck Links, in: IEEE International Conference on Communications (ICC), Dresden, Germany, 2009.
[36] X. Zhang, A. Charny, Performance Evaluation of Pre-Congestion Notification, in: International Workshop on Quality of Service (IWQoS), Enschede, The Netherlands, 2008.
[37] M. Menth, Efficiency of PCN-Based Network Admission Control with Flow Termination, Praxis der Informationsverarbeitung und Kommunikation (PIK) 30 (2) (2007) 82 – 87.
[38] M. Menth, M. Hartmann, Threshold Configuration and Routing Optimization for PCN-Based Resilient Admission Control, Computer Networks 53 (11) (2009) 1771 – 1783.
[39] J. Jiang, R. Jain, A Simple Analytical Model of Pre-Congestion Notification, in: currently under submission, 2008.
[40] D. Satoh, H. Ueno, M. Menth, Performance Evaluation of Termination in CL-Algorithm, http://tools.ietf.org/html/draft-satoh-pcn-performance-termination (Jul. 2009).
[41] P. Eardley, Traffic Matrix Scenario, http://www.ietf.org/mail-archive/web/pcn/current/msg00831.html (Oct. 2007).
[42] ISO, ISO 10589: Intermediate System to Intermediate System Routing Exchange Protocol for Use in Conjunction with the Protocol for Providing the Connectionless-Mode Network Service (1992/2002).