

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: August 08, 2012

G. Karagiannis
University of Twente
T. Taylor
Huawei
K. Chan
Consultant
M. Menth
University of Tuebingen
P. Eardley
BT
February 08, 2012

Requirements for Signaling of (Pre-) Congestion Information in a
DiffServ Domain
draft-ietf-pcn-signaling-requirements-08

Abstract

Precongestion notification (PCN) is a means for protecting quality of service for inelastic traffic admitted to a Diffserv domain. The overall PCN architecture is described in [RFC 5559](#). This memo describes the requirements for the signaling applied within the PCN domain: (1) PCN-feedback-information is carried from the PCN-egress-node to the decision point; (2) the decision point may ask the PCN-ingress-node to measure, and report back, the rate of sent PCN-traffic between that PCN-ingress-node and PCN-egress-node. The decision point may be either collocated with the PCN-ingress-node or a centralized node (in the first case, (2) is not required). The signaling requirements pertain in particular to two edge behaviors, "controlled load (CL)" and "single marking (SM)".

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months

and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 08, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](http://tools.ietf.org/html/rfc2119) [[RFC2119](http://tools.ietf.org/html/rfc2119)].

Table of Contents

1.	Introduction	3
2.	Signaling Requirements for Messages from the PCN-Egress-Nodes to Decision Point(s)	3
3.	Signaling Requirements for Messages between Decision Point(s) and PCN-Ingress-Nodes	5
4.	Security Considerations	5
5.	IANA Considerations	6
6.	Acknowledgments	6
7.	References	6
7.1.	Normative References	6
7.2.	Informative References	6
	Authors' Addresses	7

1. Introduction

The main objective of Pre-Congestion Notification (PCN) is to support the quality of service (QoS) of inelastic flows within a Diffserv domain in a simple, scalable, and robust fashion. Two mechanisms are used: admission control and flow termination. Admission control is used to decide whether to admit or block a new flow request while flow termination is used in abnormal circumstances to decide whether to terminate some of the existing flows. To support these two features, the overall rate of PCN-traffic is metered on every link in the domain, and PCN-packets are appropriately marked when certain configured rates are exceeded. These configured rates are below the rate of the link thus providing notification to boundary nodes about overloads before any congestion occurs (hence "pre-congestion" notification). The PCN-egress-nodes measure the rates of differently marked PCN traffic in periodic intervals and report these rates to the decision points for admission control and flow termination, based on which they take their decisions. The decision points may be collocated with the PCN-ingress-nodes or their function may be implemented in a centralized node.

For more details see [[RFC5559](#)],
[[draft-ietf-pcn-cl-edge-behaviour-11](#)],
[[draft-ietf-pcn-sm-edge-behaviour-08](#)].

This memo specifies the requirements on signaling protocols:

- o to carry reports from a PCN-egress-node to the decision point,
- o to carry requests, from the decision point to a PCN-ingress-node, that trigger the PCN-ingress-node to measure the PCN-sent-rate,

- o to carry reports, from a PCN-ingress-node to the decision point.

The latter two messages are only needed if the decision point and PCN-ingress-node are not collocated.

2. Signaling Requirements for Messages from the PCN-Egress-Nodes to Decision Point(s)

The PCN-egress-node measures per ingress-egress-aggregate the rates of differently marked PCN-traffic in regular intervals. The measurement intervals are recommended to take a fixed value between 100 ms and 500 ms, see [[draft-ietf-pcn-cl-edge-behaviour-11](#)], [[draft-ietf-pcn-sm-edge-behaviour-08](#)]. At the end of each measurement interval, the PCN-egress-node calculates the congestion-level-estimate (CLE) based on these quantities.

The PCN-egress-node MAY be configured to record a set of identifiers of PCN-flows for which it received excess-traffic-marked packets during the last measurement interval. The latter may be useful to perform flow termination in networks with multipath routing.

At the end of each measurement interval, or less frequently if "optional report suppression" is activated, see

[[draft-ietf-pcn-cl-edge-behaviour-11](#)], [[draft-ietf-pcn-sm-edge-behaviour-08](#)], the PCN-egress-node sends a report to the decision point.

For the SM edge behavior, the report MUST contain:

- o identifier of the PCN-ingress-node and the identifier of the PCN-egress-node (typically their IP addresses); together they specify the ingress-egress-aggregate to which the report refers,
- o rate of not-marked PCN-traffic (NM-rate) in octets/second,
- o rate of PCN-marked traffic (PM-rate) in octets/second,

For the CL edge behavior, the report MUST contain:

- o identifier of the PCN-ingress-node and the identifier of the PCN-egress-node (typically their IP addresses); together they specify the ingress-egress-aggregate to which the report refers,
- o rate of not-marked PCN-traffic (NM-rate) in octets/second,
- o rate of threshold-marked PCN traffic (ThM-rate) in octets/second,
- o rate of excess-traffic-marked traffic (ETM-rate) in octets/second,

The number format and the rate units used by the signaling protocol will limit the maximum rate that PCN can use. If signaling space is tight it might be reasonable to impose a limit, but any such limit may impose unnecessary constraints in future.

The signaling report can either be sent directly to the decision point or it can "piggy-back", i.e., be included within some other message that passes through the PCN-egress-node and then reaches the decision point.

As described in [[draft-ietf-pcn-cl-edge-behaviour-11](#)], PCN reports from the PCN-egress-node to the decision point may contain flow identifiers for individual flows within an ingress-egress-aggregate that have recently experienced excess-marking. Hence, the PCN report messages used by the PCN CL edge behavior MUST be capable of carrying sequences of octet strings constituting such identifiers."

Signaling messages SHOULD have a higher priority and a lower drop precedence than PCN-packets, see [[RFC5559](#)], to deliver them quickly and to avoid that they are dropped in case of overload.

The load generated by the signaling protocol SHOULD be minimized. We give three examples that may help to achieve that goal:

- o piggy-backing the reports by the PCN-egress-nodes to the decision point(s) onto other signaling messages that are already in place,
- o reducing the amount of reports to be sent by optional report suppression,
- o combining reports for different ingress-egress-aggregates in a single message (if they are for the same decision point).

As PCN reports are sent regularly, additional reliability mechanisms are not needed. This also holds in the presence of optional report suppression as reports are sent periodically if actions by the decision point(s) are needed, see [[draft-ietf-pcn-cl-edge-behaviour-11](#)], [[draft-ietf-pcn-sm-edge-behaviour-08](#)].

3. Signaling Requirements for Messages between Decision Point(s) and PCN-Ingress-Nodes

Through request-response signaling between the decision point and PCN-ingress-node, the decision point requests and in response the PCN-ingress-node measures and reports the PCN-sent-rate for a

specific ingress-egress-aggregate. Signaling is needed only if the decision point and PCN-ingress-node are not collocated.

The request MUST contain:

- o the identifier of the PCN-ingress-node and the identifier of the PCN-egress-node; together they determine the ingress-egress-aggregate for which the PCN-sent-rate is requested,
- o the identifier of the decision point that requests the PCN-sent-rate.

The report MUST contain:

- o the PCN-sent-rate in octets/second,
- o the identifier of the PCN-ingress-node and the identifier of the PCN-egress-node.

The request MUST be addressed to the PCN-ingress-node, and the report MUST be addressed to the decision point that requested it.

The request and the report SHOULD be sent with high priority, a lower drop precedence than PCN-packets, and reliably, because they are sent only when flow termination is needed, which is an urgent action.

Note that a complete system description for a PCN domain with centralized Decision Point includes the signaling from Decision Point to the PCN-ingress-nodes to control flow admission and termination. However, this is a known problem whose solutions were given by, for example, [[RFC3084](#)] or [[RFC5431](#)], and it lies outside the scope of the present document.

[4.](#) Security Considerations

[RFC5559] provides a general description of the security considerations for PCN. This memo relies on the security related requirements on the PCN signaling, provided in [[RFC5559](#)]. In particular, the signaling between the PCN-boundary-nodes must be protected from attacks. For example, the recipient needs to validate that the message is indeed from the node that claims to have sent it. Possible measures include digest authentication and protection against replay and man-in-the-middle attacks.

For the generic aggregate RSVP protocol, specifically, additional protection methods against security attacks are described in [[RFC4860](#)].

[5.](#) IANA Considerations

This memo includes no request to IANA.

[6.](#) Acknowledgments

We would like to acknowledge the members of the PCN working group for the discussions that produced the contents of this memo.

[7.](#) References

[7.1.](#) Normative References

- [RFC2119] S. Bradner, "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC5559] P., Eardley, "Pre-Congestion Notification (PCN) Architecture", [RFC 5559](#), June 2009.
- [[draft-ietf-pcn-cl-edge-behaviour-11](#)] T. Taylor, A. Charny, F. Huang, G. Karagiannis, M. Menth, "PCN Boundary Node Behaviour for the Controlled Load (CL) Mode of Operation (Work in progress)", December 2011.
- [[draft-ietf-pcn-sm-edge-behaviour-08](#)] A. Charny, J. Zhang, G. Karagiannis, M. Menth, T. Taylor, "PCN Boundary Node Behaviour for the Single Marking (SM) Mode of Operation (Work in progress)", December 2011.

[7.2.](#) Informative References

- [RFC3084] K. Chan, J. Seligson, D. Durham, S. Gai, K. McCloghrie, S. Herzog, F. Reichmeyer, R. Yavatkar, A. Smith, "COPS Usage for Policy Provisioning (COPS-PR)", [RFC 3084](#), March 2001.
- [RFC4860] F. Le Faucheur, B. Davie, P. Bose, C. Christou, M. Davenport, "Generic Aggregate Resource ReSerVation Protocol (RSVP) Reservations", [RFC 4860](#), May 2007.
- [RFC5431] D. Sun, "Diameter ITU-T Rw Policy Enforcement Interface Application", [RFC 5431](#), March 2009.

Authors' Addresses

Georgios Karagiannis
University of Twente
P.O. Box 217
7500 AE Enschede,
The Netherlands
EMail: g.karagiannis@utwente.nl

Tom Taylor
Huawei Technologies
1852 Lorraine Ave.
Ottawa, Ontario K1H 6Z8
Canada
Phone: +1 613 680 2675
Email: tom.taylor.stds@gmail.com

Kwok Ho Chan
Consultant

Email: khchan.work@gmail.com

Michael Menth
University of Tuebingen
Department of Computer Science
Chair of Communication Networks
Sand 13
72076 Tuebingen
Germany
Phone: +49 7071 29 70505
Email: menth@informatik.uni-tuebingen.de

Philip Eardley
BT
B54/77, Sirius House Adastral Park Martlesham Heath
Ipswich, Suffolk IP5 3RE
United Kingdom
EMail: philip.eardley@bt.com

