Internet Engineering Task Force   G. Karagiannis
Internet-Draft   University of Twente
Intended status: Informational   T. Taylor
Expires: September 02, 2011   K. Chan
                                                      Huawei Technologies
                                                                M. Menth
                                                  University of Tuebingen
                                                           March 02, 2011

**Requirements for Signaling of (Pre-) Congestion Information in a
DiffServ Domain
draft-ietf-pcn-signaling-requirements-02**

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 is carried from the PCN-egress-node to the
   decision point;(2) the decision point may demand for the measurement
   and delivery of the PCN rate sent at the PCN-ingress-node. The
   decision point may be either collocated with the PCN-ingress-node or
   a centralized node.

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 September 02, 2011.

Copyright Notice

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](RFC 2119) [[RFC2119](RFC2119)].

Table of Contents

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 as so-called PCN feedback 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-08],
   [draft-ietf-pcn-sm-edge-behaviour-05].

   This memo specifies the requirements that have to be satisfied by the
   signaling protocols needed to transport:

   o PCN egress feedback, from a PCN-egress-node to the decision point;
   o a request, from the decision point to a PCN-ingress-node, that
     triggers the PCN-ingress-node to measure the PCN-sent-rate;

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

A signaling message may either be sent directly, or may piggy-backed
on some other message that is being sent via the relevant node.

1.1. Terminology

In addition to the terms defined in [RFC5559], [draft-ietf-pcn-cl-
edge-behaviour-08] and [draft-ietf-pcn-sm-edge-behaviour-05],
this document uses the following terms:

PCN egress feedback:

   A report sent by the PCN-egress-node to the decision point. It
   reports measurements made by the PCN-egress-node that inform
   decisions about flow admission and flow termination.

PCN ingress feedback:

   A report sent by the PCN-ingress-node to the decision point. It
   reports:
     o measurements made by the PCN-ingress-node that inform
       decisions about flow termination;
     o measurements of the PCN-sent-rate.

2. Signaling requirements between PCN-egress-nodes and Decision Point

The PCN-egress-node measures the rates of differently marked PCN
traffic in regular intervals and signals them as PCN egress feedback
to the decision point.
This section describes the PCN egress feedback and the requirements
that apply to signaling protocols used for the transport of PCN
feedback from PCN-egress-nodes to decision points.
Note that if the decision point and the PCN-ingress-node are
collocated, then the signaling requirements described in this section
apply to the signaling between PCN-egress-nodes and PCN-ingress-
nodes.

2.1 Reporting Frequency

The specification of PCN-based admission control and flow termination
in [draft-ietf-pcn-cl-edge-behaviour-08], [draft-ietf-pcn-sm-edge-
behaviour-04] suggest measurement and reporting intervals at the PCN-
egress-nodes of 100 to 500 ms. The PCN reporting frequency can

provide some level of reliability. Therefore, it is considered that
for regularly reported information, additional reliability mechanisms
are not needed, see Section 2.3.6.

The following PCN contents are
sent regularly: rate of not-marked PCN traffic, rate of threshold-
marked PCN traffic, rate of excess-traffic-marked PCN traffic, CLE.

2.2 Reporting Information

This section briefly describes the information that is reported by
the PCN-egress-node.

2.2.1 PCN egress Feedback

The PCN-egress-node measures per ingress-egress-aggregate the
following rates
   o rate of not-marked PCN traffic;
   o rate of threshold-marked PCN traffic, which applies to CL edge
     behavior only;
   o rate of excess-traffic-marked PCN traffic;
   o Congestion level estimate (CLE)

The rate values are reported in octets/second to the decision point
each time that the PCN-egress-node calculates them and when this is
supported via configuration.
A report may either be sent periodically, every time the PCN-egress-
node measures the various rates, or else may be sent occasionally,
see "optional report suppression", for instance in
[draft-ietf-pcn-cl-edge-behaviour-08],
[draft-ietf-pcn-sm-edge-behaviour-05].

If so configured (e.g., because multipath routing is being used), the
PCN-egress-node MUST also include in the PCN feedback and report the
set of flow identifiers of PCN-flows for which excess-traffic-marking
was observed in the most recent measurement interval.

The representation of a flow ID depends on the surrounding
environment, e.g., "pure IP", MPLS, GMPLS, etc. Examples of such flow
ID representations can be found in [RFC2205], [RFC3175] [RFC3209],
[RFC3473].

For more details see [draft-ietf-pcn-cl-edge-behaviour-08], [draft-
ietf-pcn-sm-edge-behaviour-04].

2.3 Signaling Requirements

This section describes the requirements for signaling protocols that
are used to carry the PCN egress feedback from PCN-egress-nodes to

the decision point.

2.3.1 Priority of Signaling Messages

Signaling messages SHOULD have a higher priority than data packets.
This is needed to avoid as much as possible the situations that
during severe overload cases the signaling messages are dropped
within the PCN domain.

2.3.2 Local Information Exchange

Signaling messages MUST be able to carry the PCN egress feedback from
the PCN-egress-node to the decision point.

2.3.3 Carry Identification of PCN edge Nodes

The signaling protocol MUST carry the identity of the PCN-egress-node
that sends the message.

2.3.4 Carry Identification of ingress-egress-aggregates

The signaling protocol MUST carry the identity (address information)
of the ingress-egress-aggregates.

2.3.5 Signaling Load

The load generated by the signaling protocol SHOULD be minimized.

2.3.6 Reliability

There are situations that messages need to be received in a
reliable way. There are different ways of achieving reliability. The
specification of a mandatory solution of achieving this reliability
is out of the scope of this document. It can be however considered,
that when information is received on a regular fashion, additional
reliability measures Should Not be required.

2.3.7 Security

The PCN architecture [RFC5559] considers that all PCN-nodes are PCN-
enabled and trusted to operate correctly. In the context of this
document:
   o PCN-signaling messages MUST NOT leak out of the PCN-domain. This
     can be easily accomplished, since messages are sent to the PCN-
     boundary-node's address;
   o PCN-boundary-nodes MUST validate the signaling messages, to
     avoid that they come from an attacker. Considering that all PCN-

nodes are trusted, see [RFC5559], this requirement could be
easily fulfilled by verifying whether a message arrives on an
interface internal to the PCN-domain.

2.4. Filter Specifications

   In PCN the PCN-ingress-node and PCN-egress-nodes should be able to
   identify the ingress-egress-aggregate to which each flow belongs.
   Moreover, the PCN-egress-node also needs to associate an aggregate
   with the address of the PCN-ingress-node for receiving reports, if
   the PCN-ingress-node is the decision point. The filter specification
   at the PCN-egress-nodes depends on the surrounding environment, e.g.,
   pure IP, MPLS, GMPLS.
   In this document, a possible IP filter spec for pure IP is given as
   an example. In this case the filter spec should be able to identify a
   flow using (all or a subset of the) following information:

   o   source IP address;

   o   destination IP address;

   o   protocol identifier and higher layer (port) addressing;

   o   flow label (typical for IPv6);

   o   SPI field for IPsec encapsulated data;

   o   DSCP/TOS field;

   o   IP address of PCN-ingress-node;

   o   IP address of PCN-egress-node

3. Signaling Requirements between Decision Point and PCN-ingress-nodes

   The decision point monitors and uses the PCN egress feedback sent by
   the PCN-egress-node. There are situations that the decision point
   must obtain an estimate of the rate at which PCN-traffic is being
   admitted to the aggregate from the PCN-ingress-node.
   In order to receive this information the decision point has to
   request from the PCN-ingress-node to report the value of the PCN
   traffic admitted to a certain ingress-egress-aggregate.
   Note that if the decision point and the PCN-ingress-node are
   collocated, then the information exchanges between the decision point
   and PCN-ingress-node are internal operations.

3.1 Reporting Frequency

   The PCN content sent by the PCN-ingress-node and Decision Point are
   not sent regularly.

3.2 **Reporting Information**

3.2.1 **PCN ingress Feedback**

   The PCN-ingress-node measures per ingress-egress-aggregate the
   following rate
      o rate of admitted PCN traffic

   This value is reported in octets/second to the decision point as
   soon as possible after receiving the request from the decision
   point. This information is not sent regularly and SHOULD be
   delivered reliably.

3.2.2 **Decision Point Trigger**

   The decision point requests from the PCN-ingress-node to send for a
   certain ingress-egress-aggregate, the value of the admitted PCN
   traffic rate. This Decision Point Trigger message identifies the
   ingress-egress-aggregate for which the admitted PCN traffic rate is
   required. Moreover, since this Decision Point Trigger message sent by
   the decision point to the PCN-ingress-node is not sent regularly, it
   SHOULD be delivered reliably.

3.3 **Signaling Requirements**

   This section describes the requirements for signaling protocols that
   are used to carry the PCN ingress feedback and the Decision Point
   Trigger.

3.3.1 **Priority of Signaling Messages**

   Signaling messages SHOULD have a higher priority than data packets.

3.3.2 **Local Information Exchange**

   Signaling messages MUST be able to carry:
      o the PCN ingress feedback from the PCN-ingress-node to the
        decision point;
      o the Decision Point Trigger from the decision point to the PCN-
        ingress-node.

3.3.3 **Carry Identification of PCN edge Nodes and Decision Point**

   The signaling protocol MUST carry the identity:
      o of the PCN-ingress-node that sends the message,
      o of the decision point that sends the message.

3.3.4 **Carry Identification of ingress-egress-aggregates**

The signaling protocol MUST carry the identity (address information)
of the ingress-egress-aggregates.

3.3.5 **Signaling Load**

   The load generated by the signaling protocol SHOULD be minimized.

3.3.6 **Reliability**

   The PCN ingress feedback and the Decision Point Trigger are not sent
   regularly and SHOULD be delivered reliably. There are different ways
   of achieving reliability. The specification of a mandatory solution
   of achieving this reliability is out of the scope of this document.

3.3.7 **Security**

   The PCN architecture [RFC5559] considers that all PCN-nodes are PCN-
   enabled and trusted to operate correctly. The decision point is a
   PCN-node and therefore it is considered to be PCN-enabled and trusted
   to operate correctly. In the context of this
   document:
     o PCN-signaling messages MUST NOT leak out of the PCN-domain. This
       can be easily accomplished, since messages are sent to either the
       PCN-boundary-node's address or the decision point's address,
     o PCN-boundary-nodes MUST validate the signaling messages, to
       avoid that they come from an attacker. Considering that all PCN-
       nodes are trusted, see [RFC5559], this requirement could be
       easily fulfilled by verifying whether a message either arrives on
       an interface internal to the PCN-domain or that it is sent by a
       decision point.

4. **Security Considerations**

   [RFC5559] provides a general description of the security
   considerations for PCN. This memo introduces the additional security
   considerations described in Section 2.3.7 and Section 3.3.7.

5. **IANA Considerations**

   This memo includes no request to IANA.

6. **Acknowledgements**

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

7. References

7.1. Normative References

   [RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
               Requirement Levels", BCP 14, RFC 2119, March 1997.

   [RFC5559]   Eardley, P., "Pre-Congestion Notification (PCN)
               Architecture", RFC 5559, June 2009.

   [draft-ietf-pcn-cl-edge-behaviour-08] 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 2010.

   [draft-ietf-pcn-sm-edge-behaviour-05] 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 2010.

7.2. Informative References

   [RFC2205]   Braden, B., Zhang, L., Berson, S., Herzog, S., and S.
               Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1
               Functional Specification", RFC 2205, September 1997.
   [RFC3175]   Baker, F., Iturralde, C. Le Faucher, F., Davie, B.,
               "Aggregation of RSVP for IPv4 and IPv6 Reservations",
               RFC 3175, 2001.

   [RFC3209]   Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,
               and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP
               Tunnels", RFC 3209, December 2001.

   [RFC3473]   Berger, L., "Generalized Multi-Protocol Label Switching
               (GMPLS) Signaling Resource ReserVation Protocol-Traffic
               Engineering (RSVP-TE) Extensions", RFC 3473,
               January 2003.

Authors' Addresses

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

   Tom Taylor
   Huawei Technologies
   1852 Lorraine Ave.
   Ottawa, Ontario K1H 6Z8
   Canada
   Phone: +1 613 680 2675
   Email: tom111.taylor@bell.net

   Kwok Ho Chan
   Huawei Technologies
   125 Nagog Park
   Acton, MA 01720
   USA
   Email: khchan@huawei.com

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