Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: April 19, 2010

G. Karagiannis
University of Twente
T. Taylor
K. Chan
Huawei Technologies
M. Menth
University of Wurzburg

October 19, 2009

Requirements for Signaling of (Pre-) Congestion Information in a DiffServ Domain
draft-karagiannis-pcn-signaling-requirements-00

Status of this Memo

This Internet-Draft will expire on April 19, 2010.

Copyright Notice

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, to carry PCN content from the PCN-egress-node towards either the PCN-ingress-node or towards the centralised decision point.

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 [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 level of marking allows boundary nodes to make decisions about whether to admit or terminate. For more details see [RFC5559].

Signaling is needed to transport PCN admission control and flow termination related PCN content from PCN-egress-nodes towards either PCN-ingress-nodes or a centralised decision point, see [RFC5559]. This memo briefly describes this PCN content and it specifies the requirements that have to be satisfied by the signaling protocol needed to transport this PCN content.

Currently, only the CL [draft-ietf-pcn-cl-edge-behaviour-00] and SM [draft-ietf-pcn-sm-edge-behaviour-00] PCN edge behaviour drafts are PCN working group drafts. Therefore, the current version of this memo is only referring to the signaling requirements imposed by these drafts. If other PCN edge behaviour drafts, e.g., [draft-karagiannis-pcn-hose-edge-behaviour-00], [I-D.Piggybacked-Edge-Behaviour] will become PCN working group drafts, then a new version of this memo will also incorporate the signaling requirements imposed by these new PCN edge behaviour drafts.

1.1. Terminology

In addition to the terms defined in [RFC5559], this document uses the following terms:

Tbd.

2. Requirements for signaling from PCN-egress-nodes to PCN-ingress-nodes

This section describes the PCN information and the requirements that apply to signaling protocols used for the transport of PCN content from PCN-egress-nodes to PCN-ingress-nodes.

2.1 PCN Reporting Frequency

For the CL [draft-ietf-pcn-cl-edge-behaviour-00] and SM [draft-ietf-pcn-sm-edge-behaviour-00] the content described in the next section is reported at regular intervals, as new measurements become available. An interval of the order of 100 to 300 ms is suggested in the edge behaviour drafts.

2.2 PCN Content requirements

This section describes the PCN content, i.e., PCN information, that has to be transported by a signaling protocol from a PCN-egress-node to a PCN-ingress-node. Different types of content can be distinguished depending on the used PCN edge behaviour in use and on whether the PCN content is used during admission control or flow termination. It is important to note that the description of the PCN contents in the CL and SM edge behaviors is not yet stable. This means that the PCN contents associated with Cl and SM edge behaviours draft might change. Future versions of this memo will encompass these changes.

2.2.1 Admission control

This subsection describes which PCN contents are required to be transported from the PCN-egress-node to the PCN-ingress-node during PCN admission control. Furthermore, for each PCN content, it specifies which edge behaviours is using it.

2.2.1.1 Admission Control state

The CL [draft-ietf-pcn-cl-edge-behaviour-00] and SM [draft-ietf-pcn-sm-edge-behaviour-00] edge behaviours need to transmit a computed admission state to the point where the flow decision is made. The "Admission control state" PCN content is a boolean that uses the following values:

- admit (Boolean TRUE); this PCN content value is used to notify the PCN-ingress-node that the admission control process at the PCN-ingress-node should continue.

- block (Boolean FALSE); this PCN content value is used to notify the PCN-ingress-node that the admission control

process at the PCN-ingress-node should stop.

This PCN content can be reported by the PCN-egress-node using one of the following ways:

1) report all periodically generated PCN content (at the end of each measurement interval)
2) report only when PCN content changes, unless either ThM packets (for CL) or ETM packets (for SM) are observed. Then all periodically (per each measurement interval) must be reported.

2.2.2 Flow Termination

This subsection describes which PCN contents are required to be transported from the PCN-egress-node to the PCN-ingress-node during PCN flow termination. Furthermore, for each PCN content, it specifies which edge behaviour is using it.

2.2.2.1 Traffic rate

The CL and SM edge behaviours both need to send a traffic rate to the decision point when flow termination may be required. This rate is calculated in both cases as the number of octets per second of PCN traffic carried in packets that are not excess-marked. The CL edge behaviour calls this the estimated edge-to-edge supportable rate, while the SM edge behaviour calls it the sustainable admission rate. The processing of this rate at the decision point differs between the two edge behaviours. According to the CL edge behaviour, this rate is required (flow termination may be necessary) only when the PCN-egress-node has observed excess-marked packets in the ingress-egress aggregate being reported. The SM edge behaviour requires reporting of the traffic rate whenever the admission control state is "block".

2.2.2.2 List with flow IDs

In the case where Equal Cost Multipath (ECMP) routing is being used, if flow termination may be necessary, the CL, provide a list of flow identifiers (e.g., IP five-tuples) to the decision point. These flow identifiers indicate flows which are candidates for termination because excess-marked packets have been received within those flows. The representation of a flow ID depends on the surrounding environment, e.g., "pure IP", MPLS, GMPLS, etc. For the representation of a flow ID in a "pure IP" surrounding environment, see Section 2.3. This "List with flow IDs" PCN content can be sent when the PCN-egress-node is operating in flow termination:

- o either regularly at the end of each measurement interval;
- o or when the list, compared to previous measurement intervals, is

being modified.

2.3 Signaling requirements

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

2.3.1 General signaling requirements

This section describes the signaling requirements that are valid for both admission control and flow termination features.

2.3.1.1 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.

2.3.1.2 Local information exchange

Signaling messages MUST be able to carry the PCN contents from the PCN-egress-node to the PCN-ingress-node.

2.3.1.3 Carry identification of PCN edge nodes

The signaling protocol MUST be able to carry identification (address information) of the PCN edge nodes. However, the identification of the PCN edge nodes MUST NOT be visible outside the PCN domain.

2.3.1.4 Signaling load

The load generated by the signaling protocol to carry the PCN content from the PCN-egress-nodes to the PCN-ingress-node SHOULD be minimized as much as possible.

2.3.2 Admission control signaling requirements

This subsection describes the signaling requirements for admission control purposes.

2.3.2.1 Reliability

There are situations that PCN contents need to be sent in a reliable way, meaning that the PCN-egress-node MUST be acknowledged that the sent PCN content is successfully received by the PCN-ingress-node. It is considered that the PCN contents that are sent in a regular fashion do not need to be sent reliably.

The signaling requirements associated with each PCN content that is sent by the PCN-egress-node during admission control are described below:

- "Admission control state": The signaling protocol MUST be able to carry this PCN content, which MAY be carried reliably from the PCN-egress-node to the PCN-ingress-node.

2.3.3 Flow Termination signaling requirements

This subsection describes the signaling requirements for flow termination purposes.

2.3.3.1 Reliability

This section describes which PCN contents used during flow termination have to be sent reliably:

- "Traffic rate": The signaling protocol MUST be able to carry this PCN content, which MAY be carried reliably from the PCN-egress-node to the PCN-ingress-node.

- "List with flow IDs": The signaling protocol SHOULD be able to carry this PCN content. Moreover, this PCN contents MUST be sent reliably.

2.4. Filter specifications

The filter specification at the PCN-egress-nodes depends on the surrounding environment, e.g., pure IP, MPLS, GMPLS. In this document, only the pure IP filter spec is given as an example. In this case the filter spec should be able to identify a flow using (all of a subset of the) following information:

- source IP address;
- destination IP address;
- protocol identifier and higher layer (port) addressing;
- flow label (typical for IPv6);
- SPI field for IPsec encapsulated data;
- DSCP/TOS field.
- IP address of PCN-ingress-node

3. Requirements for PCN-egress-node to centralised decision point signaling

This section describes the PCN information and the requirements that apply to signaling protocols used for the transport of PCN content from PCN-egress-nodes to centralised decision points.

3.1 PCN Reporting Frequency

The reporting frequeny required for this type of scenario is similar to the one described in Section 2.1. The only difference is the fact that these PCN contents need to be sent from the PCN-egress-node to the centralised decision point.

3.2 PCN Content requirements

This section describes the PCN content, i.e., PCN information, that has to be transported by a signaling protocol from a PCN-egress-node to a centralized decision point. Different types of content can be distinguished depending on the PCN edge behaviour used and on whether the PCN content is used during admission control or flow termination.

3.2.1 Admission control

The same PCN contents and the same method of transmission described in Section 2.2.1 applies for this case. The only difference is the fact that these PCN contents need to be sent from the PCN-egress-node to the centralised decision pont.

3.2.2 Flow Termination

The same PCN contents and the same method of transmission described in Section 2.2.2 applies for this case. The only difference is the fact that these PCN contents need to be sent from the PCN-egress-node to the centralised decision point.

3.3 Signaling requirements

This section describes the requirements for signaling protocols that are used to carry the PCN content from PCN-egress-nodes to centralized decision points.

3.3.1 General signaling requirements

The general signaling requirements specified in Section 2.3.1 apply also for this case. The following general signaling requirements are different.

3.3.1.1 Local information exchange

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

3.3.1.2 Carry identification of PCN edge nodes

The signaling protocol MUST be able to carry identification (address information) of the PCN edge nodes and centralised decision point. However, the identification of the PCN edge nodes and the centralised decision points MUST NOT be visible outside the PCN domain.

3.3.1.3 Signaling load

The load generated by the signaling protocol to carry the PCN content from the PCN-egress-nodes to the centralized decision point SHOULD be minimized as much as possible.

3.3.2 Admission control signaling requirements

The same admission control signaling requirements described in Section 2.3.2 apply for this case. The only difference is the fact that these signaling requirements apply for signaling messages that have to be sent from a PCN-egress-node to a centralised decision point.

3.3.3 Flow Termination signaling requirements

The same flow termination signaling requirements described in Section 2.3.3 apply for this case. The only difference is the fact that these signaling requirements apply for signaling messages that have to be sent from a PCN-egress-node to a centralised decision point.

3.4. Filter specifications

The filter specification at the PCN-egress-nodes depends on the surrounding environment, e.g., pure IP, MPLS, GMPLS. The filter specifications at a PCN-egress-node described in Section 2.4 apply also for this case. The main difference is the fact that the filter specification, in this case, should be able to identify in addition to the set of parameters listed in Section 2.4, also the "IP address of the centralised decision point".

4. Security Considerations

[RFC5559] provides a general description of the security considerations for PCN. This memo introduces no new considerations.

5. IANA Considerations

This memo includes no request to IANA.

6. Acknowledgements

Tbd.

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-karagiannis-pcn-hose-edge-behaviour-00] G. Karagiannis, L. Westberg, G. Apostolopoulos, "PCN Boundary Node Behaviour for the HOSE Mode of Operation (Work in progress)", October 2009.

[draft-ietf-pcn-cl-edge-behaviour-00] 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)", July 2009.

[draft-ietf-pcn-sm-edge-behaviour-00] 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)", July 2009.

7.2. Informative References

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 Wurzburg
Institute of Computer Science
Room B206
Am Hubland, Wuerzburg D-97074
Germany
Email: menth@informatik.uni-wuerzburg.de