Internet Engineering Task Force	A. Charny
Internet-Draft	Cisco Systems
Intended status: Informational	F. Huang
Expires: March 8, 2011	Huawei Technologies
	G. Karagiannis
	U. Twente
	M. Menth
	University of Wuerzburg
	T. Taylor, Ed.
	Huawei Technologies
	September 4, 2010

TOC

PCN Boundary Node Behaviour for the Controlled Load (CL) Mode of Operation
draft-ietf-pcn-cl-edge-behaviour-07

Abstract

Pre-congestion notification (PCN) is a means for protecting the quality of service for inelastic traffic admitted to a Diffserv domain. The overall PCN architecture is described in RFC 5559. This memo is one of a series describing possible boundary node behaviours for a PCN-domain. The behaviour described here is that for a form of measurement-based load control using three PCN marking states, not-marked, threshold-marked, and excess-traffic-marked. This behaviour is known informally as the Controlled Load (CL) PCN-boundary-node behaviour.

Status of this Memo

Table of Contents

---

1. Introduction

TOC

The objective of Pre-Congestion Notification (PCN) is to protect the quality of service of inelastic flows within a Diffserv domain, in a simple, scalable, and robust fashion. Two mechanisms are used: admission control, to decide whether to admit or block a new flow request, and (in abnormal circumstances) flow termination to decide whether to terminate some of the existing flows. To achieve this, the overall rate of PCN-traffic is metered on every link in the PCN-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 PCN-boundary-nodes about incipient overloads before any congestion occurs (hence the "pre" part of pre-congestion notification). The level of marking allows decisions to be made on whether to admit or terminate PCN-flows. For more details see [RFC5559] (Eardley, P., “Pre-Congestion Notification (PCN) Architecture,” June 2009.).

PCN-boundary-node behaviours specify a detailed set of algorithms and procedures used to implement the PCN mechanisms. Since the algorithms depend on specific metering and marking behaviour at the interior nodes, it is also necessary to specify the assumptions made about PCN-interior-node behaviour. Finally, because PCN uses DSCP values to carry its markings, a specification of PCN-boundary-node behaviour must include the per domain behaviour (PDB) template specified in [RFC3086] (Nichols, K. and B. Carpenter, “Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification,” April 2001.), filled out with the appropriate content. The present document accomplishes these tasks for the controlled load (CL) mode of operation.

1.1. Terminology

TOC

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 RFC2119 [RFC2119].

In addition to the terms defined in [RFC5559] (Eardley, P., “Pre-Congestion Notification (PCN) Architecture,” June 2009.), this document uses the following terms:

> **Decision Point** The node that makes the decision about which flows to admit and to terminate. In a given network deployment, this may be the PCN-ingress-node or a centralized control node. Regardless of the location of the Decision Point, the ingress node is the point where the decisions are enforced.
>
> **NM-rate** The rate of not-marked PCN-traffic received at a PCN-egress-node for a given ingress-egress-aggregate in octets per second. For further details see Section 3.2.1 (Data Collection).

ThM-rate
The rate of threshold-marked PCN-traffic received at a PCN-egress-node for a given ingress-egress-aggregate in octets per second. For further details see Section 3.2.1 (Data Collection).

ETM-rate The rate of excess-traffic-marked PCN-traffic received at a PCN-egress-node for a given ingress-egress-aggregate in octets per second. For further details see Section 3.2.1 (Data Collection).

PCN-sent-rate The rate of PCN-traffic received at a PCN-ingress-node and destined for a given ingress-egress-aggregate in octets per second. For further details see Section 3.4 (Behaviour of the Ingress Node).

Congestion level estimate (CLE) A value derived from the measurement of PCN packets received at a PCN-egress-node for a given ingress-egress-aggregate, representing the ratio of marked to total PCN-traffic (measured in octets) over a short period. The CLE is used to derive the PCN-admission-state (Section 3.3.1 (Flow Admission)) and also by the report suppression procedure (Section 3.2.3 (Optional Report Suppression)) if report suppression is activated.

PCN-admission-state The state ("admit" or "block") derived by the Decision Point for a given ingress-egress-aggregate based on PCN packet marking statistics. The Decision Point decides to admit or block new flows offered to the aggregate based on the current value of the PCN-admission-state. For further details see Section 3.3.1 (Flow Admission).

Sustainable aggregate rate (SAR) The estimated maximum rate of PCN-traffic that can be admitted to a given ingress-egress-aggregate at a given moment without risking degradation of quality of service for the admitted flows. The intention is that if the PCN-sent-rate of every ingress-egress-aggregate passing through a given link is limited to its sustainable aggregate rate, the total rate of PCN-traffic flowing through the link will be limited to the PCN-supportable-rate for that link. An estimate of the sustainable aggregate rate for a given ingress-egress-aggregate is derived as part of the flow termination procedure, and is used to determine how much PCN-traffic must be terminated. For further details see Section 3.3.2 (Flow Termination).

CLE-reporting-threshold A configurable value against which the CLE is compared as part of the report suppression procedure. For further details, see Section 3.2.3 (Optional Report Suppression).

CLE-limit A configurable value against which the CLE is compared in order to derive the PCN-admission-state for a given ingress-egress-aggregate. For further details, see Section 3.3.1 (Flow Admission).

T-meas An interval, the value of which is configurable, defining the measurement period at the PCN-egress-node during which

statistics relating to PCN-traffic marking are collected. At the end of the interval the values NM-rate, ThM-rate, and ETM-rate as defined above are calculated and a report is sent to the Decision Point, subject to the operation of the report suppression feature. For further details see Section 3.2 (Behaviour of the PCN-Egress-Node).

T-maxsuppress An interval, the value of which is configurable, after which the PCN-egress-node must send a report to the Decision Point for a given ingress-egress-aggregate regardless of the most recent values of the CLE. This is used as a keep-alive mechanism for signalling between the PCN-egress-node and the Decision Point when report suppression is activated. For further details, see Section 3.2.3 (Optional Report Suppression).

T-fail An interval, the value of which is configurable, after which the Decision Point concludes that communication from a given PCN-egress-node has failed if it has received no reports from the PCN-egress-node during that interval. For further details see Section 3.3.3 (Decision Point Action For Missing PCN-Boundary-Node Reports).

---

2. Assumed Core Network Behaviour for CL

TOC

This section describes the assumed behaviour for nodes of the PCN-domain when acting in their role as PCN-interior-nodes. The CL mode of operation assumes that:

- PCN-interior-nodes perform threshold-marking and excess-traffic-marking of packets according to the rules specified in [RFC5670] (Eardley, P., "Metering and Marking Behaviour of PCN-Nodes," November 2009.), and any additional rules specified in the applicable encoding extension document;

- encoding of PCN status within individual packets is based on [RFC5696] (Moncaster, T., Briscoe, B., and M. Menth, "Baseline Encoding and Transport of Pre-Congestion Information," November 2009.), extended to provide a third PCN encoding state. A possible extension is documented in [ID.PCN3in1] (Briscoe, B., "PCN 3-State Encoding Extension in a single DSCP (Work in progress)," July 2010.);

- the PCN-domain satisfies the conditions specified in the applicable encoding extension document;

- on each link the reference rate for the threshold-meter is configured to be equal to the PCN-admissible-rate for the link;

- on each link the reference rate for the excess traffic meter is configured to be equal to the PCN-supportable-rate for the link;

According to [RFC5696] (Moncaster, T., Briscoe, B., and M. Menth, "Baseline Encoding and Transport of Pre-Congestion Information,"

November 2009.), the encoding extension documents should specify the allowable transitions between marking states. However, to be absolutely clear, these allowable transitions are specified here. At any interior node, the only permitted transitions are these:

- a PCN-packet that is not-marked (NM) MAY be threshold-marked (ThM) or excess-traffic-marked (ETM);

- a PCN-packet that is threshold-marked (ThM) MAY be excess-traffic-marked (ETM).

An interior node MUST NOT perform any of the following:

- re-mark a packet from PCN to non-PCN, or from non-PCN to PCN;

- re-mark a PCN-packet from threshold-marked (ThM) to not-marked (NM);

- re-mark a PCN-packet from excess-traffic-marked (ETM) to not-marked (NM) or threshold-marked (ThM).

---

3. Node Behaviours

TOC

3.1. Overview

TOC

This section describes the behaviour of the PCN-ingress-node, PCN-egress-node, and the Decision Point (which may be collocated with the PCN-ingress-node).

The PCN-egress-node collects the rates of not-marked, threshold-marked, and excess-traffic-marked PCN-traffic for each ingress-egress-aggregate and reports them to the Decision Point. It may also identify PCN-flows that have experienced excess-traffic-marking. For a detailed description, see Section 3.2 (Behaviour of the PCN-Egress-Node).

The PCN-ingress-node enforces flow admission and termination decisions. It also reports the rate of PCN-traffic sent to a given ingress-egress-aggregate when requested by the Decision Point. For details, see Section 3.4 (Behaviour of the Ingress Node).

Finally, the Decision Point makes flow admission decisions and selects flows to terminate based on the information provided by the PCN-ingress-node and PCN-egress-node for a given ingress-egress-aggregate. For details, see Section 3.3 (Behaviour at the Decision Point).

3.2. Behaviour of the PCN-Egress-Node

TOC

3.2.1. Data Collection

TOC

The PCN-egress-node MUST meter received PCN-traffic in order to derive periodically the following rates for each ingress-egress-aggregate passing through it:

- NM-rate: octets per second of PCN-traffic in PCN-packets that are not-marked;
- ThM-rate: octets per second of PCN-traffic in PCN-packets that are threshold-marked;
- ETM-rate: octets per second of PCN-traffic in PCN-packets that are excess-traffic-marked.

It is RECOMMENDED that the measurement interval, T-meas, between successive calculations of these quantities be in the range of 100 to 500 ms to provide a reasonable tradeoff between signalling demands on the network and the time taken to react to impending congestion.

The PCN-traffic SHOULD be metered continuously and the intervals themselves SHOULD be of equal length, to minimize the statistical variance introduced by the measurement process itself. The starting and ending times of the measurement intervals for different ingress-egress-aggregates MAY be the same or MAY be different.

As a configurable option, the PCN-egress-node MAY record flow identifiers of the PCN-flows for which excess-traffic-marked packets have been observed. These can be used by the Decision Point when it selects flows for termination.

> In networks using multipath routing it is possible that congestion is not occurring on all paths carrying a given ingress-egress-aggregate. Assuming that specific PCN-flows are routed via specific paths, identifying the PCN-flows that are experiencing excess-traffic-marking helps to avoid termination of PCN-flows not contributing to congestion.

3.2.2. Reporting the PCN Data

TOC

If the report suppression option described in the next sub-section is not activated, the PCN-egress-node MUST report the latest values of NM-rate, ThM-rate, and ETM-rate to the Decision Point each time that it calculates them.

If so configured (e.g., because multipath routing is being used, as explained in the previous section), the PCN-egress-node MUST also report the set of flow identifiers of PCN-flows for which excess-traffic-marking was observed in the most recent measurement interval. If this set is large, the PCN-egress-node MAY report only the most recently excess-traffic-marked PCN-flows rather than the complete set.

3.2.3. Optional Report Suppression

TOC

Report suppression MUST be provided as a configurable option, along with two configurable parameters, the CLE-reporting-threshold and the maximum report suppression interval T-maxsuppress. The default value of the CLE-reporting-threshold is zero. T-maxsuppress is discussed further at the end of this sub-section, but functions as a keep-alive mechanism for signalling between the PCN-egress-node and the Decision Point.

If the report suppression option is enabled, the PCN-egress-node MUST apply the following procedure to decide whether to send a report to the Decision Point, rather than sending a report automatically at the end of each measurement interval.

1. As well as the quantities NM-rate, ThM-rate, and ETM-rate, the PCN-egress-node MUST calculate the congestion level estimate (CLE) for each measurement interval. The CLE is equal to the ratio:

   (ThM-rate + ETM-rate) / (NM-rate + ThM-rate + ETM-rate)

   if any PCN-traffic was observed, or zero otherwise.

2. If the calculated CLE for the latest measurement interval or for the immediately previous interval is greater than the CLE-reporting-threshold, then the PCN-egress-node MUST send a report to the Decision Point. The contents of the report are described below.

3. If an interval T-maxsuppress has elapsed since the last report was sent to the Decision Point, then the PCN-egress-node MUST send a report to the Decision Point regardless of the CLE value.

4. If neither of the preceding conditions holds, the PCN-egress-node MUST NOT send a report for the latest measurement interval.

Each report sent to the Decision Point when report suppression has been activated MUST contain the values of NM-rate, ThM-rate, ETM-rate, and CLE that were calculated for the most recent measurement interval. If so configured, the PCN-egress-node MUST also report the set of flow identifiers of PCN-flows for which excess-traffic-marking was observed in the most recent measurement interval.

The above procedure ensures that at least one report is sent per interval (T-maxsuppress + T-meas). This provides some protection against loss of egress reports and also demonstrates to the Decision Point that both the PCN-egress-node and the communication path between the two nodes are in operation. However, depending on the transport used for reporting, the operator may choose to set T-maxsuppress to an effectively infinite value. For example, the transport may include its own keep-alive signalling at a frequency such that PCN keep-alive signalling is redundant.

3.3. Behaviour at the Decision Point

TOC

Operators may choose to use PCN procedures just for flow admission, or just for flow termination, or for both. The Decision Point MUST implement both mechanisms, but configurable options MUST be provided to activate or deactivate PCN-based flow admission and flow termination independently of each other at a given Decision Point.

If PCN-based flow termination is enabled but PCN-based flow admission is not, flow termination operates as specified in this document. Logically, some other system of flow admission control must be in operation, but the description of such a system is out of scope of this document and depends on local arrangements.

3.3.1. Flow Admission

TOC

The Decision Point determines the PCN-admission-state for a given ingress-egress-aggregate each time it receives a report from the egress node. It makes this determination on the basis of the congestion level estimate (CLE). If the CLE is provided in the egress node report, the Decision Point SHOULD use the reported value. If the CLE was not provided in the report, the Decision Point MUST calculate it based on the other values provided in the report, using the formula

> CLE = (ThM-rate + ETM-rate) / (NM-rate + ThM-rate + ETM-rate)

if any PCN-traffic was observed, or CLE = 0 if all the rates are zero.

The Decision Point MUST compare the reported or calculated CLE to a configurable value, the CLE-limit. If the CLE is less than the CLE-limit, the PCN-admission-state for that aggregate MUST be set to "admit"; otherwise it MUST be set to "block".

> The outcome of the comparison is not very sensitive to the value of the CLE-limit in practice, because when threshold-marking occurs it tends to persist long enough that threshold-marked traffic becomes a large proportion of the received traffic in a given interval.

If the PCN-admission-state for a given ingress-egress-aggregate is "admit", the Decision Point SHOULD allow new flows to be admitted to that aggregate. If the PCN-admission-state for a given ingress-egress-aggregate is "block", the Decision Point SHOULD NOT allow new flows to be admitted to that aggregate. These actions MAY be modified by policy in specific cases, but such policy intervention risks defeating the purpose of using PCN.

3.3.2. Flow Termination

TOC

When the report from the PCN-egress-node includes a non-zero value of the ETM-rate for some ingress-egress-aggregate, the Decision Point MUST request the PCN-ingress-node to provide an estimate of the rate (PCN-sent-rate) at which the PCN-ingress-node is receiving PCN-traffic that is destined for the given ingress-egress-aggregate.

> If the Decision Point is collocated with the PCN-ingress-node, the request and response are internal operations.

The Decision Point MUST then wait for both the requested rate from the PCN-ingress-node and the next report from the PCN-egress-node for the ingress-egress-aggregate concerned. If the next egress node report also includes a non-zero value for the ETM-rate, the Decision Point MUST determine an amount of flow to terminate using the following steps:

1. The sustainable aggregate rate (SAR) for the given ingress-egress-aggregate is estimated by the sum:

   ```
   SAR = NM-rate + ThM-rate
   ```

   for the latest reported interval.

2. The amount of traffic that should be terminated is the difference:

   ```
   PCN-sent-rate - SAR,
   ```

   where PCN-sent-rate is the value provided by the PCN-ingress-node.

If the difference calculated in the second step is positive, the Decision Point SHOULD select PCN-flows to terminate, until it determines that the PCN-traffic admission rate will no longer be greater than the estimated sustainable aggregate rate. If the Decision Point knows the bandwidth required by individual PCN-flows (e.g., from resource signalling used to establish the flows), it MAY choose to complete its selection of PCN-flows to terminate in a single round of decisions.

Alternatively, the Decision Point MAY spread flow termination over multiple rounds to avoid over-termination. If this is done, it is RECOMMENDED that enough time elapse between successive rounds of termination to allow the effects of previous rounds to be reflected

in the measurements upon which the termination decisions are based (see [IEEE-Satoh] (Satoh, D. and H. Ueno, ""Cause and Countermeasure of Overtermination for PCN-Based Flow Termination", Proceedings of IEEE Symposium on Computers and Communications (ISCC '10), pp. 155-161, Riccione, Italy," June 2010.) and sections 4.2 and 4.3 of [Menth08-sub-9] (Menth, M. and F. Lehrieder, "PCN-Based Measured Rate Termination," July 2009.)).

If the egress node has supplied a list of PCN-flow identifiers (Section 3.2 (Behaviour of the PCN-Egress-Node)), the Decision Point SHOULD first consider terminating PCN-flows in that list. In general, the selection of flows for termination MAY be guided by policy.

3.3.3. Decision Point Action For Missing PCN-Boundary-Node Reports

TOC

If the Decision Point fails to receive any report from a given PCN-egress-node for a configurable interval T-fail, it SHOULD raise an alarm to management. A Decision Point collocated with a PCN-ingress-node SHOULD cease to admit PCN-flows to the ingress-egress-aggregate passing from the PCN-ingress-node to the given PCN-egress-node, until it again receives a report from that node. A centralized Decision Point MAY cease to admit PCN-flows to all ingress-egress-aggregates destined to the PCN-egress-node concerned, until it again receives a report from that node.

If a centralized Decision Point fails to receive a reply within a reasonable period of time to a request for a PCN-sent-rate value sent to a given PCN-ingress-node, it SHOULD raise an alarm to management.

3.4. Behaviour of the Ingress Node

TOC

The PCN-ingress-node MUST provide the estimated rate of PCN-traffic received at that node and destined for a given ingress-egress-aggregate in octets per second (the PCN-sent-rate) when the Decision Point requests it. The way this rate estimate is derived is a matter of implementation.

> For example, the rate that the PCN-ingress-node supplies MAY be based on a quick sample taken at the time the information is required. It is RECOMMENDED that such a sample be based on observation of at least 30 PCN-packets to achieve reasonable statistical reliability.

3.5. Summary of Timers

TOC

Table 1 (Timers Used For the CL Boundary Node Behaviour) summarizes the timers implied by the preceding procedures. The three limits T-meas, T-maxsuppress, and T-fail apply to the three timers t-meas, t-maxsuppress, and t-fail respectively. t-meas and t-maxsuppress are reset upon expiry. t-fail is reset by management action or by receipt of a report from the PCN-egress-node concerned.

Timer	Location	Incidence	Limit	Action on Expiry
t-meas	Egress node	One per node	T-meas	Calculate and possibly report NM-rate, ThM-rate, ETM-rate and conditionally CLE for each IEA.
-	-	-	-	-
t-maxsuppress	Egress node	One per IEA if report suppression is enabled.	T-maxsuppress	Send a report for that IEA at the next expiry of T-meas.
-	-	-	-	-
t-fail	Decision point	One per egress node	T-fail	Assume failure and cease to admit flows passing through that egress node.

IEA = ingress-egress-aggregate

Table 1: Timers Used For the CL Boundary Node Behaviour

The value of T-meas SHOULD be configurable, and is RECOMMENDED to be of the order of 100 to 500 ms.

t-maxsuppress is active only when report suppression is enabled. The value of T-maxsuppress SHOULD be configurable. The appropriate value depends on the transport used to carry the egress node reports. For unreliable transport, T-maxsuppress is RECOMMENDED to be of the order of one second.

The value of T-fail MUST be configurable. When unreliable transport is used, the value of T-fail is RECOMMENDED to be of the order of 3 * T-maxsuppress if report suppression is enabled, and of the order of 3 * T-meas if report suppression is not enabled. When reliable transport is used, the operator may choose to provide similar values for T-fail or may choose to disable report timing by setting an effectively infinite value for T-fail.

4. Identifying Ingress and Egress Nodes for PCN Traffic

TOC

The operation of PCN depends on the ability of the PCN-ingress-node to identify the ingress-egress-aggregate to which each new PCN-flow belongs and the ability of the egress node to identify the ingress-egress-aggregate to which each received PCN-packet belongs. If the Decision Point is collocated with the PCN-ingress-node, the PCN-egress-node also needs to associate each ingress-egress-aggregate with the address of the PCN-ingress-node to which it must send its reports.

The means by which this is done depends on the packet routing technology in use in the network. The procedure to provide the required information is out of the scope of this document.

5. Specification of Diffserv Per-Domain Behaviour

TOC

This section provides the specification required by [RFC3086] (Nichols, K. and B. Carpenter, “Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification,” April 2001.) for a per-domain behaviour.

5.1. Applicability

TOC

This section draws heavily upon points made in the PCN architecture document, [RFC5559] (Eardley, P., “Pre-Congestion Notification (PCN) Architecture,” June 2009.).

The PCN CL boundary node behaviour specified in this document is applicable to inelastic traffic (particularly video and voice) where quality of service for admitted flows is protected primarily by admission control at the ingress to the domain. In exceptional circumstances (e.g. due to network failures) already-admitted flows may be terminated to protect the quality of service of the remaining flows. The CL boundary node behaviour is less likely to terminate too many flows under such circumstances than the SM boundary node behaviour ([I-D.SM-edge-behaviour] (Charny, A., Zhang, J., Karagiannis, G., Menth, M., and T. Taylor, “PCN Boundary Node Behaviour for the Single Marking (SM) Mode of Operation (Work in progress),” June 2010.)).

TOC

5.2. Technical Specification

The technical specification of the PCN CL per domain behaviour is provided by the contents of [RFC5559] (Eardley, P., “Pre-Congestion Notification (PCN) Architecture,” June 2009.), [RFC5696] (Moncaster, T., Briscoe, B., and M. Menth, “Baseline Encoding and Transport of Pre-Congestion Information,” November 2009.), [RFC5670] (Eardley, P., “Metering and Marking Behaviour of PCN-Nodes,” November 2009.), the specification of the encoding extension (e.g., [ID.PCN3in1] (Briscoe, B., “PCN 3-State Encoding Extension in a single DSCP (Work in progress),” July 2010.)), and the present document.

5.3. Attributes

TOC

The purpose of this per-domain behaviour is to achieve low loss and jitter for the target class of traffic. The design requirement for PCN was that recovery from overloads through the use of flow termination should happen within 1-3 seconds. PCN probably performs better than that.

5.4. Parameters

TOC

In the list that follows, note that most PCN-ingress-nodes are also PCN-egress-nodes, and vice versa. Furthermore, the PCN-ingress-nodes may be collocated with Decision Points.

Parameters at the PCN-ingress-node:

- *Filters for distinguishing PCN from non-PCN inbound traffic.
- *The markings to be applied to PCN-traffic.
- *Reference rates on each inward link for the PCN-threshold-rate and PCN-excess-rate; see Section 2 (Assumed Core Network Behaviour for CL).
- *The information needed to distinguish PCN-traffic belonging to a given ingress-egress-aggregate.

Parameters at the PCN-egress-node:

- *The measurement interval T-meas.
- *Whether report suppression is enabled and, if so, the values of the CLE-reporting-threshold and T-maxsuppress.
- *Whether individual flow identifiers must be reported for excess-traffic-marked PCN-traffic.

*The information needed to distinguish PCN-traffic belonging to a given ingress-egress-aggregate.

*The marking rules for re-marking PCN-traffic leaving the PCN domain.

Parameters at each interior node:
---

*Reference rates on each link for the PCN-threshold-rate and PCN-excess-rate; see Section 2 (Assumed Core Network Behaviour for CL).

*The markings to be applied to PCN-traffic, including the identification of PCN-packets and the encodings to indicate threshold-marking and excess-traffic-marking..

Parameters at the Decision Point:
---

*Activation/deactivation of PCN-based flow admission.

*Activation/deactivation of PCN-based flow termination.

*The value of CLE-limit.

*The maximum interval T-fail between reports from a given PCN-egress-node, for detecting failure of communications with that node.

*The information needed to map between each ingress-egress-aggregate and the corresponding PCN-ingress-node and PCN-egress-node.

5.5. Assumptions

TOC

Assumed that a specific portion of link capacity has been reserved for PCN-traffic.

5.6. Example Uses

TOC

The PCN CL behaviour may be used to carry real-time traffic, particularly voice and video.

TOC

5.7. Environmental Concerns

The PCN CL per-domain behaviour may interfere with the use of end-to-end ECN due to reuse of ECN bits for PCN marking. See the applicable PCN marking specifications for details.

5.8. Security Considerations

TOC

Please see the security considerations in Section 6 (Security Considerations) as well as those in [RFC2474] (Nichols, K., Blake, S., Baker, F., and D. Black, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers,” December 1998.) and [RFC2475] (Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, “An Architecture for Differentiated Services,” December 1998.).

---

6. Security Considerations

TOC

[RFC5559] (Eardley, P., “Pre-Congestion Notification (PCN) Architecture,” June 2009.) provides a general description of the security considerations for PCN. This memo introduces no new considerations.

---

7. IANA Considerations

TOC

This memo includes no request to IANA.

---

8. Acknowledgements

TOC

The content of this memo bears a family resemblance to [ID.briscoe-CL] (Briscoe, B., “An edge-to-edge Deployment Model for Pre-Congestion Notification: Admission Control over a DiffServ Region (expired Internet Draft),” 2006.). The authors of that document were Bob Briscoe, Philip Eardley, and Dave Songhurst of BT, Anna Charny and Francois Le Faucheur of Cisco, Jozef Babiarz, Kwok Ho Chan, and Stephen Dudley of Nortel, Giorgios Karagiannis of U. Twente and Ericsson, and Attila Bader and Lars Westberg of Ericsson.

Ruediger Geib, Philip Eardley, and Bob Briscoe have helped to shape the present document with their comments. Toby Moncaster gave a careful review to get it into shape for Working Group Last Call.

Amongst the authors, Michael Menth deserves special mention for his constant and careful attention to both the technical content of this document and the manner in which it was expressed.

9. References

TOC

9.1. Normative References

TOC

[RFC2119]	Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).
[RFC2474]	Nichols, K., Blake, S., Baker, F., and D. Black, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers,” RFC 2474, December 1998 (TXT, HTML, XML).
[RFC2475]	Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, “An Architecture for Differentiated Services,” RFC 2475, December 1998 (TXT, HTML, XML).
[RFC5559]	Eardley, P., “Pre-Congestion Notification (PCN) Architecture,” RFC 5559, June 2009 (TXT).
[RFC5670]	Eardley, P., “Metering and Marking Behaviour of PCN-Nodes,” RFC 5670, November 2009 (TXT).
[RFC5696]	Moncaster, T., Briscoe, B., and M. Menth, “Baseline Encoding and Transport of Pre-Congestion Information,” RFC 5696, November 2009 (TXT).

9.2. Informative References

TOC

[I-D.SM-edge-behaviour]	Charny, A., Zhang, J., Karagiannis, G., Menth, M., and T. Taylor, “PCN Boundary Node Behaviour for the Single Marking (SM) Mode of Operation (Work in progress),” June 2010.
[ID.PCN3in1]	Briscoe, B., “PCN 3-State Encoding Extension in a single DSCP (Work in progress),” July 2010.
[ID.briscoe-CL]	Briscoe, B., “An edge-to-edge Deployment Model for Pre-Congestion Notification: Admission Control over a DiffServ Region (expired Internet Draft),” 2006.
[IEEE-Satoh]	Satoh, D. and H. Ueno, “"Cause and Countermeasure of Overtermination for PCN-Based Flow Termination", Proceedings of IEEE Symposium on Computers and Communications (ISCC '10), pp. 155-161, Riccione, Italy,” June 2010.
[Menth08-sub-9]	Menth, M. and F. Lehrieder, “PCN-Based Measured Rate Termination,” July 2009 (PDF).
[RFC3086]	Nichols, K. and B. Carpenter, “Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification,” RFC 3086, April 2001 (TXT).

Authors' Addresses

TOC

	Anna Charny
	Cisco Systems
	300 Apollo Drive
	Chelmsford, MA 01824
	USA
Email:	acharny@cisco.com
	Fortune Huang
	Huawei Technologies
	Section F, Huawei Industrial Base,
	Bantian Longgang, Shenzhen 518129
	P.R. China
Phone:	+86 15013838060
Email:	fqhuang@huawei.com
	Georgios Karagiannis
	U. Twente
Phone:	
Email:	karagian@cs.utwente.nl
	Michael Menth
	University of Wuerzburg
	Am Hubland
	Wuerzburg D-97074
	Germany
Phone:	+49-931-888-6644
Email:	menth@informatik.uni-wuerzburg.de

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