PCN
Internet-Draft
Intended status: Informational
Expires: January 3, 2008

K. Chan
Nortel
G. Karagiannis
University of Twente
July 2, 2007

Pre-Congestion Notification Encoding Comparison
draft-chan-pcn-encoding-comparison-00

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on January 3, 2008.

Copyright Notice

Copyright (C) The IETF Trust (2007).

Abstract

DiffServ mechanisms have been developed to support Quality of Service (QoS). However, the level of assurance that can be provided with DiffServ without substantial over-provisioning is limited. Pre-Congestion Notification (PCN) investigates the use of per-flow admission control to provide the required service guarantees for the admitted traffic. While admission control will protect the QoS under normal operating conditions, an additional flow termination mechanism

is necessary in the times of heavy congestion (e.g. caused by route changes due to link or node failure).

Encoding and their transport are required to carry the congestion and pre-congestion information from the congestion and pre-congestion points to the decision points. This document provides a survey of several encoding methods, using comparisons amongst them as a way to explain their strengths and weaknesses.

Table of Contents

1. Motivation and Goals

IP networks were initially designed to perform per IP packet forwarding treatment without discrimination. With the increased use of the IP network by applications with different transport functional requirement, the notion of Quality of Service (QoS) was introduced [21].

DiffServ [10] introduced differentiated per packet forwarding treatment to provide QoS: some packets are served at a higher scheduling priority than others. Diffserv Service Classes [19] categorises various DiffServ traffic and recommends how they can be used for packets from applications with different transport requirements. For instance there are Telephony and Real-time Interactive service classes. Applications like these need low loss, low delay and low jitter. A suitable Per Hop Behavior (PHB) is Expedited Forwarding (EF) [16], which works by assuring that packets (usually) encounter very short or empty queues. Each router is allocated a certain amount of bandwidth for the EF PHB, for instance. Excess packets are dropped and delayed, thus leading to poorer QoS for an end user running an application like voice-over-IP. Even if average traffic levels are known, due to traffic variations the level of assurance that can be provided with DiffServ without substantial over-provisioning is limited.

To help ensure that the average traffic loads remain within the allocated bandwidth limits, the DiffServ Architecture [10] introduces the idea of policing the amount of traffic in a class as it enters the network. The acceptable traffic level is described by a traffic

conditioning agreement (TCA). However, TCAs police the aggregate traffic in a class at the network ingress, and for scalability reasons typically includes traffic to different destinations. As a result, TCA's do not guarantee that EF aggregate at any given node in the network does not exceed the allocated capacity [23], and so don't ensure that a particular end user's QoS is guaranteed. Also, in practice TCAs are static and so require accurate and/or conservative prediction of the traffic matrix. Also, in practice the TCA at the ingress must allow any destination address, if it is to remain scalable.

To cope with the issue of exceeding bandwidth allocation to EF on some links, in practice a policer or shaper is assumed to be installed at the interior nodes as well. However, shaping or policing traffic causes excess packets be dropped and delayed, thus leading to poorer QoS for an end user running an application like voice-over-IP. Even if average traffic levels remain within the allocated bandwidth limits, traffic variations may limit the level of assurance that can be provided with DiffServ without substantial

over-provisioning.

These factors motivate us to work on per flow admission control for a DiffServ network, and in particular on measurement-based admission control, ie new flow requests are blocked dynamically in response to actual (incipient) congestion on a router within the DiffServ network.

However, despite flow admission control, sometimes there can be heavy congestion - for example caused by link or node failure that effectively reduces the network's capacity. The default option is that the QoS of all flows is degraded. However, by terminating some flows the QoS of the remaining flows can be protected. The work reported in I-D.silverman-tsvwg-mlefphb indicates that in the context where calls have different recongizable precedence levels (e.g. in the context of military/emergency calls [22]), this problem can be partially addressed by dropping lower-precednce calls preferentially while protecting higher precedence calls. However, as it was shown in [6], the need to terminate some flows of a given precedence level, while protecting the QoS of the rest of the flows of this precedence level remains.

This motivates us to work on per flow termination for a DiffServ network, and in particular on measurement-based termination, ie existing on-going flows are dropped dynamically in response to actual congestion on a router within the DiffServ network.

Explicit Congestion Notification (ECN) [15] introduced the idea of a router indicating that it is congested by changing the header of packets ("marking" them). However, ECN in RFC3168 [15] is designed for TCP applications. This motivates us to develop the concept for real-time applications. A router "PCN-marks" packets as an early warning of its incipient congestion ("pre-congestion"). These markings are then used by the admission control and termination mechanisms.

The main goal of this document is a survey and comparison of several encoding and transport methods that are required to encode the pre-congestion information and to transport it from the PCN interior nodes to the decision PCN egress nodes. In order to accomplish this comparison, a number of criteria are developed. The possible encoding and transport methods are:

- Encoding and transport using the combination of the ECN and DSCP bits of a data packet header
- Encoding and transport using the ECN bits of a data packet header

- Encoding and transport using the DSCP bits of a data packet header
- Encoding and transport using a different channel than data packets

The rest of this document is organized as follows. Section 2 describes the encoding requirements indicated by currently known detection and marking mechanisms that can be used within the PCN-domain. Section 3 describes a survey of the possible encoding and transport methods. The comparison of these methods is accomplished in Section 4 and Section 5 provides the conclusion. The rest of the sections describe the security considerations, acknowledgements, IANA considerations and references.

2. PCN Encoding Requirements and Features in current PCN Detection,

Marking, and Transport Mechanisms

In order to derive a number of encoding and transport methods it is important to be aware of which PCN based mechanisms are used for congestion and pre-congestion detection and marking. Therefore, this section describes the PCN encoding and transport features and the encoding modes/states that are possible in the current available PCN based algorithms used for congestion and pre-congestion detection and marking in interior nodes. The current PCN detection, marking and transport mechanisms are briefly introduced in the Appendix of this document and are discussed in detail in CL PHB [5], Single-Marking [3], Three-State-Marking [2] and Load-Control [4].

The main PCN features that can be supported by the PCN based algorithms introduced in the Appendix of this document are:

- "admission control", see PCN-Architecture [1]
- "flow termination", see PCN-Architecture [1]
- "not congested", used to identify/notify that a congestion and/or a pre-congestion situation has not occurred in a certain communication path.
- "ECMP handling", used to solve the ECMP problem that is related to the fact that flows that are not passing through a congested PCN interior node can belong to an aggregate that detects a congestion. Any measures that are taken on such flows will not solve the congestion problem, since such flows are not contributing and causing the congestion on the PCN interior node.

Furthermore, it is important to emphasize that in general, dealing with the ECMP handling during flow termination, could be somewhat

disjoint from how a detection and marking algorithm operates. For example:

1. The CL-PHB [5] and/or Single-Marking [3] algorithm, similar to the Load-Control [4] algorithm, could use the "Affected Marking", encoding mode/state, see Appendix A.4, to solve the ECMP problem at the expense of an additional DSCP value and the expense of keeping track of which flows have been Affected Marked and which

have not.

2. The CL-PHB [5] and/or Single-Marking [3] algorithm, similar to the Three-State-Marking [2] algorithm, could choose for termination only flows which have been Termination Marked at the expense of additional complexity at the edge of needing to keep track of which flows have been Termination Marked or not.

2.1. Supported PCN Features and Encoding States in CL-PHB Method

In CL-PHB [5], see also Appendix A.1, a solution has been developed that can be used in PCN-domains, to provide the admission control and flow termination features. Furthermore, this algorithm can support the "not congested" feature, which is used to notify that a congestion and/or a pre-congestion situation has not occurred in a certain communication path.

The algorithm currently specified in CL-PHB [5] does not specify if and how the "ECMP handling" feature is supported. Therefore, it can be deduced that currently, CL-PHB [5] supports the following main PCN supported encoding features: the "not congested", "admission control", and the "flow termination".

The congestion and precongestion information is mainly encoded and transported by using the combination of the ECN and DSCP field carried in the IP header of the data packets. The used PCN encoding and transport modes/states are:

- "Admission Marking" used by the "admission control" feature

- "Termination Marking" used by the "flow termination" feature

Due to the fact that among others, ECN bits are used to transport the congestion and pre-congestion information, the ECN-nonce modes/states have to also be transported. In particular, the ECN-nonce modes/states are used to support the "not congested" feature. Furthermore, the "Not-ECN capable" mode/state needs to be used in order to indicate to a node that the traffic is not ECN-capable. The Explicit Congestion Notification (ECN)-nonce is an optional addition to ECN that protects against accidental or malicious concealment of marked

packets from the TCP sender. It uses the two ECN-Capable Transport

(ECT) codepoints in the ECN field of the IP header. It improves the robustness of congestion control by enabling co-operative senders to prevent receivers from exploiting ECN to gain an unfair share of network bandwidth. The ECN-Capable Transport (ECT) codepoints '10' and '01' (ECT(0) and ECT(1) respectively) are set by the data sender to indicate that the end-points of the transport protocol are ECN-capable.

In particular, the main encoding scheme used in CL-PHB [5] is given by Option 1 in Figure 1.

2.2. Supported PCN Features and Encoding States in Three State PCN Marking Method

The solution proposed in Three-State-Marking [2] supports the "admission control", "flow termination", and "not congested" features. Furthermore this solution can also support the "ECMP handling" feature during the flow termination process. This feature can be provided using the explicit excess load marking approach, a marked packet belongs to a flow that was routed through congested router. Therefore an additional marking mode/state for the support of the "ECMP handling" feature is not needed.

Thus the main PCN supported encoding modes/states are:

- "Admission Marking" used by the "admission control" feature
- "Termination Marking" used by the "flow termination" and "ECMP handling" features.
- "Not congested Marking" used by the "not congested" feature.

The exact method of transporting the congestion and precongestion information is not specified in Three-State-Marking [2], but the method given by Option 1 in Figure 1 (or number of other encoding options) can be used.

2.3. Supported PCN Features and Encoding States in Single Marking Method

The solution proposed in Single-Marking [3], see also Appendix A.3, supports the "admission control" and "flow termination" and "not congested" features. The algorithm currently specified in Single-Marking [3], similar to the algorithm specified in CL-PHB [5], does not specify if and how the "ECMP handling" feature is supported.

The way of how the congestion and precongestion information is

transported is not described in Single-Marking [3], but it is emphasized that it can be similar to the transportation way used in CL-PHB [5]. As mentioned in Section 2.1, due to the fact that among others, ECN bits are used to transport the congestion and pre-congestion information, the ECN-nonce modes and Not ECN-capable mode have to also be transported. Thus the main PCN supported encoding modes/states are:

- "Admission Marking" used by the "admission control" and "flow termination" features.

A possible way of how the encoding scheme can be implemented for the Single-Marking [3] mechanism is given by Option 3 (or number of other encoding options) in Figure 1.

2.4. Supported PCN Features and Encoding States in Load Control Method

The algorithm proposed in Load-Control [4], see also Appendix A.4, supports the "admission control", "flow termination", "not congested" and "ECMP handling" features. Note that this algorithm provides solutions to the ECMP problem that can occur during either the admission control or the flow termination process.

The congestion and precongestion information is transported by using the DSCP field carried in the IP header of the data packets. Thus the main PCN supported encoding modes/states are:

- "Admission Marking" used by the "admission control" feature (Encoding Option 10, see section 3.3.1).

- "Termination (or Encoded) Marking" used by the "flow termination" feature (and in Encoding Option 11, see section 3.3.2, also used by the "admission control" feature).

- "Not congested Marking" used by the "not congested" feature.

- "Affected Marking" that in combination with the ""Termination (or Encoded) Marking" is used to support the "ECMP handling" feature.

In particular, the main encoding scheme used in Load-Control [4] is given by Option 10 and Option 11 in Figure 1. With details provided in section 3.3.1 and 3.3.2.

3. Survey of Encoding and Transport Methods

There are many choices available for encoding the PCN information. To provide a summary and an overview, we use the following table of

current proposed encodings. Clarifying the abreviation and normiclature used in the table and some description of each of these encoding choices and their trade-offs are in subsequent sub sections.

ECN Bits	00	01	10	11	DSCP
RFC 3168	Not-ECT	ECT(1)	ECT(0)	CE	NA
Option 1	AM	ECT(1)	ECT(0)	TM	PCN
Option 2	AM	ECT(A)	ECT(T)	TM	PCN
Option 3	Not-ECT	ECT(1)	ECT(0)	AM/TM	PCN
Option 4	Not-ECT	ECT(1)	ECT(0)	AM	PCN+TM
Option 5	Not-ECT	ECT(1)	ECT(0)	AM or TM	NA
Option 6	Not-ECT	ECT(A)	ECT(T)	AM/TM	NA
Option 7	AM	ECT(A)	ECT(T)	TM	NA
Option 8	Not-CE	AM	PM	NDS-CE	NA
Option 9	Not-ECT	ECT	AM	TM	NA
Option 10	NA	NA	NA	NA	4 DSCP
Option 11	NA	NA	NA	NA	3 DSCP

Figure 1: Encoding of PCN Information in IP Header

Notes for Figure 1: Options 10 and 11 use different DSCPs to encode the PCN states, hence the indication of 4 DSCPs and 3 DSCPs (for 4 PCN states and 3 PCN states respectively). The NA under the ECN bits simply means the use of the ECN bits are Not Applicable for these options. Details on the 4 DSCPs and 3 DSCPs usage are provided in

sections 3.3.1 and 3.3.2 respectively.

The above table contains abreviations of terms, their meaning are as follows:

- ECN Bits: This refers to the two bit field in the IP header defined by RFC 3168 [15].

- DSCP: DiffServ Code Point. This refers to the six bit field in the IP header defined by RFC 2474 [10].
- Not-ECT: Not ECN Capable Transport. Defined in RFC 3168 [15].
- ECT(0), ECT(1): ECN Capable Transport. Defined in RFC 3168 [15].
- CE: Congestion Experienced. Defined in RFC 3168 [15].
- NA: Not Applicable. Meaning this field is not used for this encoding choice.
- AM: Admission Marked.
- TM: Termination Marked.
- PCN: The DSCP field uses a specific code point for PCN traffic.
- Not-CE: Not experiencing congestion. This have the same meaning as ECT(0) and ECT(1), but without the cheater detection functionality.
- NDS-CE: Not DiffServ capable traffic with congestion experienced.

The encoding states/modes required are

- PCN Capable Transport Marking, for separation from None PCN Capable Transport
- Not congested Marking, for indicaton of No Congestion Indication
- Admission Marking, for indication of Flow Admission Information

- Termination Marking, for indication of Flow Termination Information
- Nonce Marking, for cheater detection
- Affected Marking for ECMP indication

The possible encoding and transport methods are:

- Encoding and transport using the combination of the ECN and DSCP bits of a data packet header
- Encoding and transport using the ECN bits of a data packet header

- Encoding and transport using the DSCP bits of a data packet header
- Encoding and transport using a different channel (e.g., IPFIX, see RFC 3955 [18]) than the IP header of the data packets

The encoding table provided in Figure 1 is organized following the general encoding method given above. With the exception of not describing the "different channel" method. Following sub-sections provide additional details to each of the Encoding Option choices. Further more, some possible use of these encoding states are summarized by the detection methods descriptions in Appendix A. But we encourage the reader to read each of the PCN detection algorithm drafts as continual improvements are made based on simulation work.

3.1. Encoding and Transport Using Both ECN and DSCP Fields

This section describes the Encoding Options that uses the combination of ECN and DSCP bits available in the IP header of data packets to encode the PCN states.

One feature of this group of Encoding Options sets them apart from the others: They all use the inherent nature of DiffServ for traffic class separation to fullfil the PCN Encoding State requirment of: PCN Capable Transport Marking. This use of DiffServ and DSCP will also satisfy the need to keep none PCN Capable traffic out of the PCN

Capable traffic class. Hence this group of Encoding Options will view the rest of the required PCN encoding states/modes as being subset of being part of PCN Capable traffic class.

Note that these encoding schemes are denoted in this document as "Encoding Option 1", "Encoding Option 2", "Encoding Option 3", and "Encoding Option 4". The transport of the congestion and pre-congestion information is accomplished using the IP data packets.

3.1.1. Option 1 Encoding

As compared to the encoding indicated by RFC 3168 [15], because the requirement for indicaton of PCN Capable traffic and None PCN Capable traffic is being handled by DSCP, the "00" bit encoding is being used for Admission Marking indication. Leaving the "11" for Termination Marking indication. The Nonce Marking and the Not Congested Marking requirement is provided by the use of ECT(1)/01 and ECT(0)/10.

Hence Encoding Option 1 statisfies PCN Encoding requirements of:

- PCN Capable Transport Marking, for separation from None PCN Capable Transport

- Not congested Marking, for indicaton of No Congestion Indication
- Admission Marking, for indication of Flow Admission Information
- Termination Marking, for indication of Flow Termination Information
- Nonce Marking, for cheater detection

With the PCN Encoding requirement not satisfied being:

- Affected Marking for ECMP indication

3.1.2. Option 2 Encoding

Encoding Option 2 builds on Encoding Option 1 and adds the additional capability of the sender specifying interest of receiving Admission Marking or Termination Marking information by using ECT(A)/01 and

ECT(T)/10. This additional control and separation of Admission and Termination information may provide the PCN edge nodes added capabilities, which are out of scope for this document.

As with Encoding Option 1, Encoding Option 2 statisfies PCN Encoding requirements of:

- PCN Capable Transport Marking, for separation from None PCN Capable Transport
- Not congested Marking, for indicaton of No Congestion Indication
- Admission Marking, for indication of Flow Admission Information
- Termination Marking, for indication of Flow Termination Information
- Nonce Marking, for cheater detection

With the PCN Encoding requirement not satisfied being:

- Affected Marking for ECMP indication

3.1.3. Option 3 Encoding

Encoding Option 3 uses a single marking to represent both Admission Information and Termination Information. This saving of a marking code point allows the restoraton of None PCN Capable Transport indicaton of Not-ECT/00. Allowing this encoding to look more like the RFC 3168 [15] encoding (in encoding syntax, encoding semantax is

not represented here). But the None PCN Capable Transport requirement is already provided for by the use of DiffServ and DSCP, hence there is no additional functional difference with Encoding Option 1 and 2.

Encoding Option 3 statisfies PCN Encoding requirements of:

- PCN Capable Transport Marking, for separation from None PCN Capable Transport
- Not congested Marking, for indicaton of No Congestion Indication

- Admission Marking, for indication of Flow Admission Information
- Termination Marking, for indication of Flow Termination Information
- Nonce Marking, for cheater detection

With the PCN Encoding requirement not satisfied being:

- Affected Marking for ECMP indication

3.1.4. Option 4 Encoding

Encoding Option 4 uses a new DSCP to indicate Termination Information. Instead of using code point within the ECN bits. This introducton of a new DSCP will require DiffServ to handle traffic marked with this new DSCP the same way as all other PCN traffic. Besides this difference, Encoding Option 4 is very much like Encoding Option 5 and RFC 3168 [15]'s encoding.

Encoding Option 4 statisfies PCN Encoding requirements of:

- PCN Capable Transport Marking, for separation from None PCN Capable Transport
- Not congested Marking, for indicaton of No Congestion Indication
- Admission Marking, for indication of Flow Admission Information
- Termination Marking, for indication of Flow Termination Information
- Nonce Marking, for cheater detection

With the PCN Encoding requirement not satisfied being:

- Affected Marking for ECMP indication

3.2. Encoding and Transport Using ECN Field

This section describes the Encoding Options that uses only the ECN bits available in the IP header of data packets to encode the PCN states.

Please notice this group of Encoding Options does not use DiffServ at all. Hence there are no separatoin of traffic based on their DSCP values and DiffServ classes. With this group of Encoding Options, the required states of "PCN Capable Transport"/"None PCN Capable Transport" must be encoded using the ECN bits. Also, without the protection/separation capability provided by DiffServ, it is typically harder to satisfy the requirement set forth in RFC 4774 [20] for alternate ECN semantics.

Note that these encoding schemes are denoted in this document as "Encoding Option 5", "Encoding Option 6", "Encoding Option 7", and "Encoding Option 8". The transport of the congestion and pre-congestion information is accomplished using the IP data packets.

3.2.1. Option 5 Encoding

Encoding Option 5 is actually identical to the encoding provided by RFC 3168 [15]. With the option of using the bit pattern of 11 to represent the AM or TM state. Encoding Option 5's simularity to RFC 3168 [15]'s encoding allows it to be easily understood by people who understands RFC 3168 [15]. But at the same time, this gives it the most difficulty when satisfying the requirements set forth in RFC 4774 [20] is needed.

The use of Not-ECT will separate PCN traffic from none PCN traffic with the big exception of for ECN traffic.

Encoding Option 5 statisfies PCN Encoding requirements of:

- Not congested Marking, for indicaton of No Congestion Indication
- Admission Marking, for indication of Flow Admission Information
- Termination Marking, for indication of Flow Termination Information
- Nonce Marking, for cheater detection

With the PCN Encoding requirement not satisfied being:

- PCN Capable Transport Marking, for separation from None PCN Capable Transport
- Affected Marking for ECMP indication

3.2.2. Option 6 Encoding

Encoding Option 6 uses the ECT(A)/01 and ECT(T)/10 Markings to indicate what kinds of information the sender wants, and hence indicates if the CE/11 Marking indicates Admission or Termination information.

But Encoding Option 6 suffers the same issue as Encoding Option 5 on ability to separate traffic and indications between PCN and ECN traffic. Hence they have the same problem satisfying the requirements set forth in RFC 4774 [20].

Encoding Option 6 statisfies PCN Encoding requirements of:

- Not congested Marking, for indicaton of No Congestion Indication
- Admission Marking, for indication of Flow Admission Information
- Termination Marking, for indication of Flow Termination Information
- Nonce Marking, for cheater detection

With the PCN Encoding requirement not satisfied being:

- PCN Capable Transport Marking, for separation from None PCN Capable Transport
- Affected Marking for ECMP indication

3.2.3. Option 7 Encoding

Encoding Option 7 sacrafies the indication of None PCN Capable Transport to allow the use of a different code point for indicating Admission information. But this still suffers the same problems as Encoding Options 5 and 6.

Encoding Option 7 statisfies PCN Encoding requirements of:

- Not congested Marking, for indicaton of No Congestion Indication
- Admission Marking, for indication of Flow Admission Information

- Termination Marking, for indication of Flow Termination Information
- Nonce Marking, for cheater detection

With the PCN Encoding requirement not satisfied being:

- PCN Capable Transport Marking, for separation from None PCN Capable Transport
- Affected Marking for ECMP indication

3.2.4. Option 8 Encoding

Encoding Option 8 gives up the ability to provide the Nonce capability for allowing the indication of RFC 3168 [15] Congestion Experienced (CE) and PCN indications at the same time. But then it can not distinguish PCN/ECN Capable traffic from None PCN/ECN Capable traffic, and still suffers the same issues as Encoding Options 5, 6, and 7.

Encoding Option 8 statisfies PCN Encoding requirements of:

- Not congested Marking, for indicaton of No Congestion Indication
- Admission Marking, for indication of Flow Admission Information
- Termination Marking, for indication of Flow Termination Information

With the PCN Encoding requirement not satisfied being:

- PCN Capable Transport Marking, for separation from None PCN Capable Transport
- Nonce Marking, for cheater detection
- Affected Marking for ECMP indication

3.2.5. Option 9 Encoding

Encoding Option 9 gives up the ability to provide the Nonce capability for allowing separate code points for Admission information and Termination information. It also retains the ability to indicate Not PCN Capable Transport. But it still suffers the lack of ability to be distinguished from RFC 3168 [15] ECN traffic.

Encoding Option 9 statisfies PCN Encoding requirements of:

- Not congested Marking, for indicaton of No Congestion Indication
- Admission Marking, for indication of Flow Admission Information
- Termination Marking, for indication of Flow Termination Information

With the PCN Encoding requirement not satisfied being:

- PCN Capable Transport Marking, for separation from None PCN Capable Transport
- Nonce Marking, for cheater detection
- Affected Marking for ECMP indication

3.3. Encoding and Transport Using DSCP Field

In this type of encoding and transport method the congestion and precongestion information is encoded into the 6 DSCP bits that are transported in the IP header of the data packets. Two possible alternatives can be distinguished, as indicated in the following sub-sections.

3.3.1. Option 10 Encoding

Each of the encoding modes/states use a separate DSCP value, meaning that when all encoding modes/states are supported then 4 DSCP values are needed for encoding. Note that all DSCP values are representing and are associated with the same PHB. The supported encoding modes/states supported by this scheme are

- DSCP0 [original value] "Not congested Marking"

- DSCP1 [first additional experimental value] "Admission Marking"

- DSCP2 [second additional experimental value] "Termination (Encoded) Marking"

- DSCP3 [third additional experimental value] "Affected marking"

3.3.2. Option 11 Encoding

Each of the "Not congested Marking", "Termination (Encoded) Marking" and "Affected marking" modes/states use a different DSCP value. Note that in this alternative the "termination (Encoded) Marking" mode/ state is used to support both the "admission control" and "flow termination" features. This means that 3 DSCP values are needed for

encoding. Note that all DSCP values are representing and are associated with the same PHB.

The supported encoding modes/states supported by this scheme are:

- DSCP0 [original value] "Not congested Marking"

- DSCP1 [first additional experimental value] "Termination (Encoded) Marking"

- DSCP2 [second additional experimental value] "Affected marking"

3.4. Encoding and Transport Using IPFIX

In this type of encoding and transport method the congestion and precongestion information can be encoded using the IPFIX protocol RFC 3955 [18], that is normally used to carry flow-based IP traffic measurements from an observation point to a collecting point. Note that this encoding scheme is denoted in this document as "IPFIX channel". An observation point is a location in a network where IP packets can be observed and measured. A collecting point can be a process or a node that receives flow records from one or more observation points. In the PCN case, each PCN-interior-node will be an IPFIX observation point and the PCN-egress-node will be the IPFIX collecting point.

The PCN-interior-node will support the metering process and the flow records. Note that in this case each flow record can be associated with the record of the congestion and pre-congestion metering information associated with each PHB. The PCN-egress-node will then support the IPFIX collecting process, which will receive flow records from one or more congested and pre-congested PCN-interior-nodes. Using this encoding method the encoding modes/states can be aggregated and transported to the egress node by using the flow records at regular intervals or at the moment that a congestion and pre-congestion situation occurs. The used transport channel in this case is not the data path but a signaling protocol.

4. Encoding Comparison

This section provides a comparison between the encoding and transport methods described in Section 3. In order to do this comparison a number of criteria are derived mainly by studying the current PCN detection, marking and transport mechanisms described in Section 2.

4.1. Comparison Criteria

The following subsections describe a number of criteria that can be used to compare the encoding and transport methods discussed in Section 3.

4.1.1. Co-Existence of PCN and Non-PCN Traffic

This criterion emphasizes whether the used mechanisms allow the coexistence of PCN traffic and of non-PCN traffic within the same PCN-domain. The non-PCN traffic represents the traffic that cannot become PCN marked and it belongs to another PHB than the PCN-traffic.

4.1.2. Supported PCN Features

This criterion is used to evaluate how many and which PCN features are supported by an encoding and transport scheme. The PCN features are:

- Not congested
- Admission control
- Flow termination
- ECMP handling

4.1.3. Required Encoding States/Modes

This criterion is used to evaluate how many and which encoding modes/ states are supported by an encoding scheme.

The possible PCN encoding modes are (note that some of them can be combined):

- Not PCN-capable: - used to indicate to a node that the traffic is not PCN- capable. By using this encoding mode a distinction can be made between PCN- traffic and non PCN-traffic, see Section 4.1.1.
- "Not congested Marking", typically used to support the "not congested" Feature
- "Admission marking", typically used to support the "admission control" Feature
- "Termination marking", typically used to support the "flow termination" feature

- "Affected Marking" used to support the "ECMP handling" feature.

When the ECN bits are used to transport the congestion and pre-congestion information, the ECN-nonce modes and the Not ECN-capable mode have to also be transported:

- ECT(1) marking
- ECT(0) marking
- Not ECN-capable - used to indicate to a node that the traffic is not ECN-capable.

Note that the ECT(1) and ECT(0) modes/states are the ECN nonce modes/ states and are used to support the "not congested" feature.

4.1.4. Encoding Implementation Requirements

This criterion emphasizes the encoding implementation requirements, regarding the need and the manner of using DSCPs, PHBs, ECN bits or other type of encoding.

4.1.5. Different ECN Semantics Capability

This criterion is representing the first alternate ECN semantics issue discussed in [RFC4774]. This criterion only applies to encoding and transport schemes that are using the alternate ECN semantics.

"(1) The first issue concerns how routers know which ECN semantics to use with which packets in the network:

How does the connection indicate to the router that its packets are using alternate ECN semantics? Is the specification of alternate-ECN semantics robust and unambiguous? If not, is this a problem?

As an example, in most of the proposals for alternate ECN semantics, a diffserv field is used to specify the use of alternate ECN semantics. Do all routers that understand this diffserv codepoint understand that it uses alternate ECN semantics, or not? Diffserv allows routers to re-mark DiffServ Code Point (DSCP) values within the network; what is the effect of this on the alternate ECN semantics?" from [RFC4774]

4.1.6. Old Router Impacts

This criterion is representing the second and third alternate ECN semantics issues discussed in [RFC4774]. This criterion only applies

to encoding and transport schemes that are using the alternate ECN semantics.

"(2) A second issue is that of incremental deployment in a network where some routers only use the default ECN semantics, and other

routers might not use ECN at all. In this document, we use the phrase "new routers" to refer to the routers that understand the alternate ECN semantics, and "old routers" to refer to routers that don't understand or aren't willing to use the alternate ECN semantics.

The possible existence of old routers raises the following question: How does the possible presence of old routers affect the performance of the alternate-ECN connections?

(3) The possible existence of old routers also raises the question of how the presence of old routers affects the coexistence of the alternate-ECN traffic with competing traffic on the path.", from [RFC4774].

4.1.7. Alternate-ECN Traffic Performance

This criterion is the fourth alternate ECN semantics issue discussed in [RFC4774]. This criterion only applies to encoding and transport schemes that are using the alternate ECN semantics.

"(4) A final issue is that of the general evaluation of the alternate ECN semantics:

How well does the alternate-ECN traffic perform, and how well does it coexist with competing traffic on the path, in a "clean" environment with new routers and with the unambiguous specification of the use of alternate ECN semantics?", from [RFC4774]

In particular, the following detailed issues should be taken into account:

- Verification of Feedback from the Router (see Section 5.1 in [RFC4774])
- Coexistence with Competing Traffic (see Section 5.2 in [RFC4774])
- Proposals for Alternate ECN with Edge-to-Edge Semantics (see Section 5.3 in [RFC4774])
- Encapsulated Packets (see Section 5.4 in [RFC4774])

- A General Evaluation of the Alternate ECN Semantics (see Section 5.5 in [RFC4774])

4.2. Encoding and Transport Comparison

This section describes the comparison of the encoding and transport methods described in section 3, by using the criteria described in Section 4.1. The encoding schemes are indicated in Figure 1.

The comparison is presented in the following way. Each subsection describes a comparison of the encoding schemes indicated in Figure 1 based on one of the criteria introduced in Section 4.1.

4.2.1. Co-Existence of PCN and Non-PCN Traffic

The Encoding Option 9 scheme is the only scheme that is allowing the coexistence of PCN and non-PCN traffic. The rest of the schemes described in Section 3 are not allowing the coexistence of PCN and non-PCN traffic. This can however, be realized when an additional encoding mode/state is used, i.e., the Not PCN-capable mode described in Section 4.2.3, which allows to distinguish between the non PCN-traffic and the PCN-traffic. This additional encoding mode/state can be realized by using DiffServ to separate the PCN traffic for all other none PCN traffic.

4.2.2. Supported PCN Features

The Encoding Option 10, Encoding Option 11, and "IPFIX channel" schemes can support the four PCN features: "not congested", "Admission control", "Flow termination" and "ECMP handling".

The Encoding Option 1, Option 6, Option 2, Option 4, and Option 5 schemes are able to support the PCN features "not congested", "admission control" and "flow termination". Furthermore, the Option 9 scheme can support the PCN features "admission control" and "flow termination" and the Option 5 can support the "not congested" "flow termination" and "ECMP handling" features.

Note that Encoding Option 1, Option 6, Option 2, Option 4, Option 9, Option 5 (AM) could also support the "ECMP handling" feature, used during the flow termination process, when the algorithm that uses these encoding modes/states could choose for termination only flows which have been Termination Marked at the expense of additional complexity at the edge of needing to keep track of which flows have been Termination Marked or not.

4.2.3. Supported Encoding States/Modes

The "IPFIX channel" solution does not use the encoding modes/states listed in Section 4.1.3. This is because the "IPFIX channel" encoding solution does not use the data path for encoding and transport, but it requires to use a separate signaling channel to transport the congestion and pre-congestion information associated with the "not congested", "admission control", "flow termination" and "ECMP handling" PCN features.

The "Not PCN-capable" encoding mode is not used by the presented encoding schemes. However, if the separation between the PCN traffic and non-PCN traffic is required, then the "Not PCN-capable" has to be used by all schemes.

The "Not congested Marking" encoding mode is used by:

- Encoding Option 10
- Encoding Option 11

The "Admission Marking" encoding mode/state is used by:

- Encoding Option 10
- Encoding Option 11
- Encoding Option 1
- Encoding Option 6
- Encoding Option 2
- Encoding Option 4
- Encoding Option 9
- Encoding Option 5

The "Termination Marking" encoding mode/state is used by:

- Encoding Option 10

- Encoding Option 1
- Encoding Option 6

- Encoding Option 2
- Encoding Option 4
- Encoding Option 9
- Encoding Option 5

The "Affected Marking" encoding mode/state is used by:

- Encoding Option 10
- Encoding Option 11

The "ECN-nonce" encoding modes ((ECT(1) and ECT(0)) marking are used by:

- Encoding Option 1
- Encoding Option 6
- Encoding Option 2
- Encoding Option 4
- Encoding Option 5A
- Encoding Option 5T

The "Not ECN-capable" encoding mode is used by:

- Encoding Option 1
- Encoding Option 6

- Encoding Option 2
- Encoding Option 4
- Encoding Option 9
- Encoding Option 5A
- Encoding Option 5T

4.2.4. Encoding Implementation Requirements

The "IPFIX channel" encoding mode needs a separate signaling channel for the transport of the congestion and precongestion information from the PCN-interior-nodes towards the PCN-egress-node. The requirement of using an additional channel increases the complexity and influences negatively the performance of the PCN-interior-nodes since each PCN-interior-node needs to support in addition to the data path a separate channel.

Encoding Option 10 and 11 (the DSCP-Alternatives) need to use in addition to the original DSCP, three DSCP and two DSCP values, respectively. These additional DSCP values can be taken from the DSCP values that are not defined by standards action, see [8]. Note that all the DSCP values are representing and are associated with one PHB. Furthermore, if the separation between the PCN traffic and non-PCN traffic is required, then an additional DSCP or PHB value is needed for the "Not PCN-capable" encoding mode. The value of this DSCP/PHB can either follow a standards action or use a value that is applied for experimental or local use. It is important to note that the number of the DSCP values used for local or experimental use is restricted.

Encoding Options 1 to 9 (the ECN-Alternatives) need to take into account, in addition to the PCN encoding modes also the encoding modes that are specific to ECN, which are the "ECN nonce" and "Not ECN-Capable" modes. Encoding Options 6, 9, 5A, 5T need to only use the 4 ECN values. The use of the ECN values has to comply to

[RFC4774], see also Section 4.2.5, 4.2.6, 4.2.7. The rest of the ECN-Alternatives, i.e., Option 1, 2, 3, 4 need to use the 4 ECN values and one DSCP value. As mentioned above, the use of the ECN values has to comply to [RFC4774], see also Section 4.2.5, 4.2.6, 4.2.7. Furthermore, the additional DSCP value can either be defined using a standard action or by using, similar to Option 10 and 11 (the DSCP-Alternatives), a DSCP value defined for experimental or local use.

Furthermore, for all ECN-Alternatives, with exception to Option 9, an additional DSCP or PHB value is needed for the encoding of the "Not PCN-capable" mode. The value of this DSCP/PHB can either follow a standards action or use a value that is applied for experimental or local use. An alternative to using another DSCP, the points of view of all traffic not DSCP marked with PCN may be considered "Not PCN-capable". This may be applicable only to Encoding Options that uses DIffServ.

4.2.5. Different ECN Semantics Capability

To satisfy the first alternate ECN semantics issue discussed in [RFC4774] on "how does the connection indicate to the router that its packets are using alternate ECN semantics?", the PCN traffic will need to be distinguishable from the none PCN traffic and other ECN traffic.

There are actually two issues indicated here. First: the ability to distinguish PCN traffic from none PCN traffic. Second: the ability to distinguish PCN traffic from ECN traffic.

For solving the first issue, the use of "Not-ECT" state to indicate none PCN (also none ECN) traffic will be sufficient. But this does not solve the second issue of distinguishing PCN traffic from ECN traffic. The use of DSCP to distinguish PCN traffic from all other traffic will solve both issues indicated.

With the use of a specific DSCP to indicate PCN traffic, encoding Option 1, Option 2, Option 3, Option 4 of Figure 1 (Encoding of PCN Information in IP Header) will satisfy this issue. The other

encoding Options will solve only one or the other issue, not solve both issues.

4.2.6. Old Router Impacts

The second issue and the third issue raised by [RFC4774] is concerned with the existence of both PCN routers and none PCN routers. The use of a PCN DSCP allows the segregation of the PCN traffic away from the other traffic. With the single PCN domain edge-to-edge deployment scenario, all devices are at least DiffServ capable and controlled by one management entity. With the use of the PCN DSCP, and correct configuration of DiffServ, these two issues are resolved.

With the use of a specific DSCP to indicate PCN traffic, encoding Option 1, Option 2, Option 3, Option 4 of Figure 1 (Encoding of PCN Information in IP Header) will satisfy this issue. The other encoding Options will solve only one or the other issue, not solve both issues.

4.2.7. Alternate-ECN Traffic Performance

The forth issue raised by [RFC4774] is related to the performance of the PCN semantics. This issue is more related to the marking algorithm using the encoding to transport the PCN information. Hence will not handle this issue until a later version of this document.

5. Conclusions

To Be Filled In After PCN List Discussions.

6. Security Implications

Packets from normal precedence and higher precedence sessions [22] aren't distinguishable by PCN Interior Nodes. This prevents an attacker specifically targeting, in the data plane, higher precedence packets (perhaps for DoS or for eavesdropping). However, PCN End Nodes can access this information to help decide whether to admit or terminate a flow. The separation of network information provided by the Interior Nodes and the precedence information at the PCN End

Nodes allows simpler, easier and better focused security enforcement.

PCN End Nodes police packets to ensure a flow sticks within its agreed limit. This is similar to the existing IntServ behaviour. Between them the PCN End Nodes must fully encircle the PCN-Region, otherwise packets could enter the PCN-Region without being subject to admission control, which would potentially destroy the QoS of existing flows.

It is assumed that all the Interior Nodes and PCN End Nodes run PCN and trust each other (ie the PCN-enabled Internet Region is a controlled environment). For instance a non-PCN router wouldn't be able to alert that it's suffering pre-congestion, which potentially would lead to too many calls being admitted (or too few being terminated). Worse, a rogue router could perform attacks such as marking all packets so that no flows were admitted.

So security requirements are focussed at specific parts of the PCN-Region:

- The PCN End Nodes become the trust points. The degree of trust required depends on the kinds of decisions it has to make and the kinds of information it needs to make them. For example when the PCN End Node needs to know the contents of the sessions for making the decisions, when the contents are highly classified, the security requirements for the PCN End Nodes involved will also need to be high.

- PCN-marking by the Interior Nodes along the packet forwarding path needs to be trusted, because the PCN End Nodes rely on this information.

7. IANA Considerations

To be completed.

8. Acknowledgements

To be completed.

Appendix A. Current PCN Detection, Marking and Transport Mechanisms

This appendix describes briefly the available PCN based mechanisms that can be used for congestion and pre-congestion detection and marking used at interior nodes. The following subsections focus on the main characteristics of such algorithms that are influencing the encoding and transport features, which are the encoding and marking modes/states and the used transport channel. The current PCN detection, marking and transport algorithms are discussed in detail in CL-PHB [5], Single-Marking [3], Three-State-Marking [2] and Load-Control [4].

Appendix A.1. Detection, Marking and Transport Mechanisms in CL-PHB

This section describes briefly the detection, marking and transport algorithm specified in CL-PHB [5]. As a fundamental building block to enable the admission control and flow termination algorithms, each link of the PCN- domain is associated with a configured-admissible-rate and configured-termination-rate; the former is usually significantly larger than the latter. If traffic in a specific DiffServ class ("PCN-traffic") on the link exceeds these rates then packets are marked with "Admission-Marking" or "Termination-Marking".

To support the admission control algorithm, each PCN-interion-node in the PCN-domain runs an algorithm to determine whether to Admission Mark the packet. The algorithm measures the PCN-traffic on the link and ensures that packets are admission marked before the actual queue builds up. The algorithm's main parameter is the configured-admissible-rate, which is set lower than the link speed. Admission marked packets indicate that the PCN traffic rate is reaching the configured-admissible-rate and so act as an "early warning" that the engineered capacity is nearly reached. Therefore they indicate that requests to admit prospective new PCN flows may need to be refused. The Admission Marked and Termination Marked packets are transported downstream towards the PCN-egress-node. The PCN-egress-node then uses the received Admission Marked and Termination Marked pakets to measure the Congestion-Level-Estimate for traffic from each remote PCN-ingress-node. The Congestion-Level-Estimate is the number of

bits in PCN packets that are Admission marked or Termination marked, divided by the number of bits in all PCN packets. It is calculated by an PCN-egress-node separately for the PCN packets from each particular PCN-ingress-node. This Congestion-Level-Estimate provides an estimate of how near the links on the path inside the PCN-domain are getting to the configured-admissible-rate. Subsequently, the Congestion-Level-Estimate is signaled to the PCN-ingress-node. The PCN-ingress-node uses the CLE value for admission control, i.e., when the CLE is higher than a threshold then new flow requests are rejected.

To support flow termination, each node in the PCN-domain runs an algorithm to determine whether to Terminate Mark the packet. The algorithm measures the PCN traffic and ensures that packets are Termination Marked before the actual queue builds up. The algorithm's main parameter is the configured-termination- rate, which is set lower than the link speed (but higher than the configured-admissible-rate). Thus termination marked packets are transported downstream towards the PCN- egress-node to indicate that the PCN traffic rate is reaching the configured- termination-rate and so act as an "early warning" that the engineered capacity is nearly reached. Therefore they indicate that it may be advisable to terminate some of the existing PCN flows in order to preserve the QoS of the others.

The PCN-egress-node calculates also per ingress-egress aggregate the Sustainable Admission Rate (SAR), which is actually the rate of the received unmarked PCN-traffic. The SAR is sent to the PCN-ingress-node that is used to calculate the amount of flows that have to be terminated in order to stop the severe congestion situation. This is accomplished by measuring, per ingress - egress aggregate, the PCN-traffic that is destined for the specific PCN-egress-node and by subtracting the SRA from it in order to calculate the excess amount of PCN flows that have to be terminated.

Appendix A.2. Detection, Marking and Transport Mechanisms in Three State Marking

Please see draft-babiarz-pcn-3sm-00.txt [2] for details on the Three State Marking Algorithm.

Appendix A.3. Detection, Marking and Transport Mechanisms in Single Marking

This section describes briefly the detection, marking and transport algorithm specified in Single-Marking [3].

The PCN-Interior-node meters the PCN traffic and marks the excess rate. It is important to note that only one single marking procedure

is needed for admission control and flow termination. The admission marking rate is proportional to the excess rate above the configured-admissible-rate. Since the rate at which admission has to be stopped is preferably significantly lower than the rate at which flow termination is required, which is the main argument for having two different markings, the single marking solution has to provide for different levels of admission and flow termination as well. To do this the solution introduces a system-wide constant u which is the ratio configured-termination-rate/configured-admissible-rate.

The PCN-egress-node measures the rate of both PCN marked and PCN unmarked traffic on a per-ingress egress aggregate basis, and reports to the PCN-ingress-node two values: the rate of PCN unmarked traffic from this PCN-ingress-node, which is denoted as Sustainable Admission Rate (SAR) and the Congestion Level Estimate (CLE), which is the fraction of the marked traffic received from this PCN-ingress-node.

The SAR is calculated by measuring the amount of received PCN unmarked rate. The Congestion Level Estimate (CLE) is calculated in a similar way as specified in CL-PHB [5]. Both values are calculated for each ingress-egress aggregate and they are reported to these PCN-ingress-nodes. Each PCN-ingress-node calculates the Sustainable Preemption Rate (SPR) by simply multiplying SAR with the system-wide constant u. The termination (or pre-emption) of flows only takes place when the rate of all flows sent by the PCN-ingress-node exceeds the SPR. The number of flows to be terminated is calculated in the following way. Per ingress - egress aggregate, the PCN-ingress-node measures the PCN- traffic that is destined for the specific PCN-egress-node and by subtracting the SPR from it in order to calculate the excess amount of PCN flows that have to be terminated.

Appendix A.4. Detection, Marking and Transport Mechanisms in Load Control Marking

This section describes briefly the detection, marking and transport algorithm specified in Load-Control [4].

This algorithm is supporting the admission control and flow termination features. The admission control feature based on probing can be used to implement a simple measurement-based admission control within a Diffserv domain. In these PCN-interior-nodes, thresholds are set for the traffic belonging to different PHBs in the measurement based admission control function. In this scenario an IP

packet is used as a probe packet, meaning that the DSCP field in the header of the IP packet is re-marked when the predefined configured admissible-rate is exceeded. When the predefined configured admissible-rate is exceeded all packets are remarked by a node. In this way also the data packets are marked to notify the PCN-egress-

node that a congestion occurred on a particular PCN-ingress-node to PCN-egress-node path. The PCN edges can then admit or reject flows that are requesting resources. The rate of the re-marked data packets is used to detect a congestion situation that can influence the admission control decisions.

By using probing, the ECMP problem that is associated with the admission control feature can be, to a certain degree, solved by being able to identify which flows are passing through the congested node.

The flow termination feature is able to terminate flows in case of exceptional events, such as severe congestion after re-routing. The exceptional event, or severe congestion can be detected using a DSCP remarking approach where the packet remarking is proportional to the amount of unavailable resources. In particular, the Diffserv nodes mark packets whenever the measured link throughput rate exceeds a configured-termination-rate and the proportion of the marked packets is in proportion to the excess traffic above the configured-termination-rate threshold. This type of marking is denoted as encoded marking and the marked packets are denoted as Encoded Marked packets. It is important to note that any data packets that are passing through the congested node and are not Encoded Marked are marked differently using another DSCP value. This type of marking is denoted as Affected Marking and the marked packets are denoted as Affected Marked packets.

The PCN-egress-nodes can use the Encoded Marked packets to calculate the percentage of throughput or bandwidth that does exceed the configured-termination-rate threshold. The PCN-egress-node can then, in combination with the PCN-ingress-node, the sender of the traffic and the support of the PCN domain(s), reduce the generated throughput, by terminating ongoing flows, until the configured-termination-rate threshold is satisfied. Note that the PCN-egress-node will select only flows that received Encoded Marked and Affected Marked data packets. In this way the ECMP problem is solved by being

able to identify which flows are passing through the congested node.

9. Informative References

[1] Eardley, P., "Pre-Congestion Notification Architecture", draft-eardley-pcn-architecture-00 (work in progress), June 2007.

[2] Babiarz, J., "Three State PCN Marking", draft-babiarz-pcn-3sm-00 (work in progress), June 2007.

[3] Charny, A., "Pre-Congestion Notification Using Single Marking for Admission and Termination", draft-charny-pcn-single-marking-02 (work in progress), July 2007.

[4] Westberg, L., "LC-PCN - The Load Control PCN solution", draft-westberg-pcn-load-control-00 (work in progress), May 2007.

[5] Briscoe, B., "Pre-Congestion Notification marking", draft-briscoe-tsvwg-cl-phb-03 (work in progress), October 2006.

[6] Baker, F. and J. Polk, "MLEF Without Capacity Admission Does Not Satisfy MLPP Requirements", draft-ietf-tsvwg-mlef-concerns-00 (work in progress), February 2005.

[7] Braden, B., Clark, D., and S. Shenker, "Integrated Services in the Internet Architecture: an Overview", RFC 1633, June 1994.

[8] Wroclawski, J., "Specification of the Controlled-Load Network Element Service", RFC 2211, September 1997.

[9] Braden, B., Clark, D., Crowcroft, J., Davie, B., Deering, S., Estrin, D., Floyd, S., Jacobson, V., Minshall, G., Partridge, C., Peterson, L., Ramakrishnan, K., Shenker, S., Wroclawski, J., and L. Zhang, "Recommendations on Queue Management and Congestion Avoidance in the Internet", RFC 2309, April 1998.

[10] 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.

[11] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, December 1998.

[12] Heinanen, J., Baker, F., Weiss, W., and J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, June 1999.

[13] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and J. McManus, "Requirements for Traffic Engineering Over MPLS", RFC 2702, September 1999.

[14] Bernet, Y., Ford, P., Yavatkar, R., Baker, F., Zhang, L., Speer, M., Braden, R., Davie, B., Wroclawski, J., and E. Felstaine, "A Framework for Integrated Services Operation over Diffserv Networks", RFC 2998, November 2000.

[15] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, September 2001.

[16] Davie, B., Charny, A., Bennet, J., Benson, K., Le Boudec, J., Courtney, W., Davari, S., Firoiu, V., and D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop Behavior)", RFC 3246, March 2002.

[17] Charny, A., Bennet, J., Benson, K., Boudec, J., Chiu, A., Courtney, W., Davari, S., Firoiu, V., Kalmanek, C., and K. Ramakrishnan, "Supplemental Information for the New Definition of the EF PHB (Expedited Forwarding Per-Hop Behavior)", RFC 3247, March 2002.

[18] Leinen, S., "Evaluation of Candidate Protocols for IP Flow Information Export (IPFIX)", RFC 3955, October 2004.

[19] Babiarz, J., Chan, K., and F. Baker, "Configuration Guidelines for DiffServ Service Classes", RFC 4594, August 2006.

[20] Floyd, S., "Specifying Alternate Semantics for the Explicit

Congestion Notification (ECN) Field", BCP 124, RFC 4774, November 2006.

[21] "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanisms", Proceedings of SIGCOMM '92 at Baltimore MD, August 1992.

[22] "Multilevel Precedence and Pre-emption Service (MLPP)", ITU-T Recommendation I.255.3, 1990.

[23] "Economics and Scalability of QoS Solutions", BT Technology Journal Vol 23 No 2, April 2005.

Authors' Addresses

Kwok Ho Chan
Nortel
600 Technology Park Drive
Billerica, MA 01821
USA

Email: khchan@nortel.com

Georgios Karagiannis
University of Twente
P.O. Box 217
7500 AE Enschede,
The Netherlands

Email: g.karagiannis@ewi.utwente.nl

Full Copyright Statement

Copyright (C) The IETF Trust (2007).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).