Abstract

This document proposes an algorithm for marking and metering by using pre-congestion notification (PCN) baseline encoding for both flow admission and flow termination. The proposed algorithm uses threshold marking whose TBthreshold.threshold is set to the number of bits of a metered-packet smaller than the token bucket size. Another threshold for the token bucket is required to change the marking. frequency. Under the flow termination state, all the packets are to be threshold-marked. Under the admission stop state, 1/Nth of packets are to be threshold-marked when the meter indicates. We
evaluates the performance of the proposed algorithm by simulations. Our simulation indicates that the performance is almost the same as that of the CL[I-D.briscoe-tsvwg-cl-architecture] algorithm.

Table of Contents

1. Introduction ............................................. 3
2. Terminology ............................................. 3
3. Single Threshold-Marking ................................. 3
   3.1. Operation at the PCN-interior-node ................. 4
   3.2. Operation at the PCN-egress-node ................. 7
   3.3. Operation at the PCN-ingress-node ................. 7
       3.3.1. Admission Decision ......................... 7
       3.3.2. Flow Termination Decision ................. 7
4. Impact on PCN marking behaviour .......................... 8
5. Acknowledgements ......................................... 8
6. Appendix A: Simulation Setup and Environment ............ 8
   6.1. Network and Signaling Models ....................... 8
   6.2. Traffic Models ..................................... 10
   6.3. Performance Metrics ................................ 12
   6.4. Parameter setting for STM .......................... 12
   6.5. Parameter setting for CL ........................... 12
   6.6. Simulation Environment .............................. 13
7. Appendix B: Admission Control Simulation .................. 13
   7.1. Parameter setting for admission control ............ 13
   7.2. Basic evaluation .................................. 13
   7.3. Effect of Ingress-Egress Aggregation ............... 14
       7.3.1. CBR ......................................... 14
       7.3.2. VBR ......................................... 15
       7.3.3. SVD ......................................... 15
   7.4. Effect of multi-bottleneck ........................... 16
   7.5. Fairness among different Ingress-Egress pairs ....... 17
8. Appendix C: Flow termination ............................... 18
   8.1. Basic evaluation .................................. 18
9. Appendix D: Another Flow Termination Control ............. 19
10. IANA Considerations ..................................... 19
11. Security Considerations .................................. 19
12. Informative References .................................. 20
Authors’ Addresses .......................................... 20
Intellectual Property and Copyright Statements ............ 22
1. Introduction

Pre-congestion notification (PCN) gives information to support admission control and flow termination in order to protect the quality of service (QoS) of inelastic flows. Although several algorithms (e.g., [I-D.briscoe-tsvwg-cl-phb], [I-D.charny-pcn-single-marking], [I-D.babiarz-pcn-3sm]) have been proposed to achieve PCN, only the single marking algorithm (SM) [I-D.charny-pcn-single-marking] meets the requirement of baseline encoding.

This document proposes an algorithm for marking and metering by using PCN baseline encoding for both flow admission and flow termination. Our algorithm uses PCN-threshold marking although SM uses PCN-excess-traffic marking. Although our algorithm uses an elaborate mechanism, it chooses PCN-admissible and PCN-supportable rates independently.

2. Terminology

The terminology used in this document conforms to the terminology of [ID.ietf-pcn-architecture] and [I-D.ietf-pcn-marking-behaviour].

3. Single Threshold-Marking

We propose an algorithm of marking and metering by using PCN baseline encoding for both flow admission and flow termination. This algorithm for marking uses only the PCN-threshold-rate as PCN-supportable rate and does not use the PCN-excess-rate. We show a schematic of how the PCN admission control and flow termination mechanisms operate as the rate of PCN-traffic increases, for a PCN-domain with three type of ratios of PCN-threshold-marked packets and three states of Marking.ratio at Fig. 1. As shown Fig. 1, no packets are PCN-marked at a rate less than the PCN-admissible-rate. Some packets are PCN-marked at rates between PCN-admissible and PCN-supportable rates, and all the packets are PCN-marked at rates greater than the PCN-supportable-rate. Single threshold-marking algorithm (STM) uses a PCN-threshold-rate with a very large TBthreshold. threshold and one-N-th packets are marked (Marking.frequency: 1/N) according to PCN-threshold marking with a very large TBthreshold.threshold to accomplish having some packets PCN-marked. Marking.frequency is a ratio of actual marking to marking according to PCN-threshold marking. One-N-th packets are regularly marked according to PCN-threshold marking with a very large TBthreshold.threshold. Marking with Marking.frequency: 1/N is similar to N-th packet marking of [I-D. westberg-pcn-load-control]. The marking.frequency is applied to threshold marking although N-th packet marking is applied to excess traffic marking. To achieve
having no packets being PCN-marked under the PCN-admissible rate, the marking switch is turned off under the PCN-admissible rate. To achieve having all the packets PCN-threshold marked, the proposed algorithm uses PCN-threshold marking with another TBthreshold. This marking.frequency is 1, and this is the same as described in [I-D. pcn-marking-behaviour] when the other TBthreshold is regarded as TBthreshold.threshold in [I-D. pcn-marking-behaviour]. We explain it in detail in the following subsections.

![Figure 1: Ratio of PCN-threshold-marked packets, Control operations, and Marking.frequency](image)

### 3.1. Operation at the PCN-interior-node

We assume that a token bucket mechanism is used; other implementations are possible. We explain how to mark by using two token buckets in cooperation. The first token bucket is not used for marking but for a marking switch. The second token bucket is used for marking. When the marking switch is OFF, no packet is marked. When the switch is ON, some or all packets are marked according to the second token bucket.

We explain the first token bucket in detail. Tokens of the first token bucket are added at the PCN-admissible-rate. Tokens are removed equal to the size in bits of the metered-packet. These
additions and removals are independent of the second token bucket. If the metered traffic is sustained at a level greater than the PCN-admissible-rate, then the marking switch is set to ON. However, a short burst of traffic at greater than the PCN-admissible-rate may not trigger the switch to go ON. If the burst is sufficiently long such that the amount of tokens in the first token bucket are less than threshold (TB1.threshold), the marking switch is set to ON. Otherwise, the marking switch is OFF. The behavior of the first token bucket is similar to that of Threshold-marking although there is a difference between actual marking and having the marking switch ON. When the marking switch is ON, packets are PCN-threshold-marked according to the meter based on the second token bucket. When the marking switch is OFF, no packet is marked regardless of the meter of the second token bucket.

We explain the second token bucket in detail. Regardless of the marking switch being ON or OFF, tokens of the second token bucket are added at the PCN threshold-rate, which is the PCN-supportable-rate and tokens are removed equal to the size in bits of the metered-packet. These additions and removals are independent of the first token bucket. The TBthreshold.threshold should be set to the number of bits of a metered packet smaller than the token bucket size although it is not an *intermediate* depth described in [I-D. ietf-pcn-marking-behaviour]. This is the same as the second token bucket size before the removal of tokens for the arrived packet. If the marking switch is ON and Marking.frequency is 1, an arrived packet is marked. If the marking switch is ON and Marking.frequency is 1/N, 1/Nth of arrived packets are marked. The Marking.frequency has two values: 1 and 1/N. If the amount of tokens in the second token bucket is not more than TB2.threshold, the Marking.frequency is 1. Otherwise, it is 1/N. The TB2.threshold is set at a deep depth of the second token bucket. The number of N MUST be the same in the domain. This PCN-domain-wide constant does not impose any constraint except the PCN-domain-wide constant itself.

Note that marking by metering whether the amount of token in the second token bucket is less than the TBthreshold.threshold or not gives a good approximation of utilization of a single link. If packets arrive according to the Poisson process, the marking ratio is the utilization of a single link [Kleinrock]. Furthermore, if an arrival process is a superposition of renewal processes, the arrival process approaches the Poisson process as the number of superpositions approaches infinity [Cox]. Therefore, when the number of superpositions of renewal processes is sufficiently large, this metering gives a good approximation of the utilization of a single link.

An arrived packet is PCN-threshold-marked per one-Nth when the
marking switch is ON and the Marking.frequency is 1/N, that is, the amount of tokens of the second token buckets is less than the TBthreshold.threshold. If the amount of tokens of the second token bucket is less than a configured intermediate depth, termed TB2.threshold, the Marking.frequency is changed 1/N to 1. The Marking.frequency is changed from 1 to 1/N when tokens of the second token bucket is larger than the TB2.threshold. This marking is expressed by the following pseudo code:

\[
\begin{align*}
\text{Input: pcn packet} \\
\text{Marking.ratio} &= 1/N; \\
n &= 0; \\
\text{FOR } i = 1 \text{ to } 2 \\
\text{TB}_i\text{.fill} &= \min(\text{TB}_i\text{.size}, \text{TB}_i\text{.fill} + \text{TB}_i\text{.rate}\times(\text{now} - \text{TB}_i\text{.lastUpdate})); \\
&\quad \text{// add tokens to the two token buckets} \\
\text{TB}_i\text{.fill} &= \min(\text{TB}_i\text{.fill} - \text{packet.size}, 0); \\
&\quad \text{// remove tokens from each token bucket} \\
\text{ENDFOR} \\
\text{IF } \text{TB}_1\text{.fill} \leq \text{TB}_1\text{.threshold} \text{ THEN } \text{// marking switch is ON} \\
\text{IF } \text{TB}_2\text{.fill} < \text{TB}_2\text{.threshold}.threshold \text{ THEN} \\
\text{IF } \text{TB}_2\text{.fill} > \text{TB}_2\text{.threshold} \text{ THEN} \\
\text{markCnt}++; \\
&\quad \text{IF } \text{mod}\text{(markCnt, N)} == 0 \text{ THEN } \text{// Marking.frequency = 1/N} \\
&\quad\quad \text{markCnt} = 0; \\
&\quad\quad \text{packet.mark} = \text{TM}; \\
&\quad \text{ENDIF} \\
\text{ELSE} \\
\text{markCnt} = 0; \\
\text{packet.mark} = \text{TM}; \\
\text{ENDIF} \\
\text{ENDIF} \\
\text{ELSE } \text{// marking switch is OFF} \\
\text{ENDIF} \\
\text{output: void}
\end{align*}
\]

where TB2.fill represents TBthreshold.fill in [I-D. pcn-marking-behaviour], TB2.size represents TBthreshold.max in [I-D. pcn-marking-behaviour], TB2.rate represents the PCN-threshold-rate (as PCN-supportable-rate), TB2.lastUpdate represents TBthreshold.lastUpdate, and TB2threshold.threshold represents TBthreshold.threshold in [I-D. pcn-marking-behaviour]. TB2.threshold represents a threshold at which the Marking.frequency from is changed from 1/N to 1 or from 1 to 1/N, as explained above, which is not defined in [I-D. pcn-marking-behaviour]. Threshold marked(TM) represents PCN-thresholdmarked according to Supportable rate.
3.2. Operation at the PCN-egress-node

A PCN-egress-node measures the ratio of PCN-threshold-marked packets on a per-ingress basis, and reports to the PCN-ingress-node one value. The value is the Congestion Level Estimate (CLE), which is the fraction of the marked traffic received from the ingress node and is exactly the same as that of CL, SM, and three state PCN marking algorithm (3sM).

Furthermore, PCN-egress-node measures L-sequential marked packets per PCN-ingress-node. When the PCN-egress-node receives those packets from a PCN-ingress-node, it starts measuring the receiving rate of this PCN-ingress-node. L-sequential marked packets received during the measurement interval are ignored. The PCN-egress-node sends a packet of information about the receiving rate to the PCN-ingress-node after the measurement interval. After that, the PCN-egress-node again starts measuring the receiving rate if it receives L-sequential marked packets.

3.3. Operation at the PCN-ingress-node

3.3.1. Admission Decision

Just as in CL and SM, the admission decision is based on the CLE. The ingress node stops admission of new flows if the CLE is greater than a predefined threshold (CLE threshold). The CLE threshold is chosen under the following maximum value. The maximum of the CLE threshold MUST be

\[ \text{CLE threshold} = \frac{1}{N} \times \rho \]

where \( \rho \) represents the minimum ratio between the PCN-admissible-rate and PCN-supportable-rate via all the links between PCN-ingress and egress nodes and \( N \) is the denominator of marking frequency \( 1/N \).

Note that the marking ratio approaches to the maximum CLE threshold described above when the load is PCN-admissible-rate at a bottleneck link via all the links between PCN-ingress and egress nodes as the number of superpositions of flows approaches infinity, as described in Sect 3.1.

3.3.2. Flow Termination Decision

When a PCN-ingress-node receives a packet with information about a receiving rate from a PCN-egress-node, it starts measuring a sending rate for this PCN-egress-node immediately. If the PCN-ingress-node receives a new receiving rate from the same PCN-egress-node during measurement interval, the new rate is ignored. At the end of the
measurement interval, the PCN-ingress node terminates flows whose bandwidth is the same as the sending rate - (1-y)* receiving rate. The value of y is small enough to tolerate over termination.

This termination spends more time when the PCN-supportable rate is much less than the physical rate because the sending rate is almost the same as the receiving rate even under a congested situation.

4. Impact on PCN marking behaviour

The goal of this section is to propose several minor changes to the PCN-marking behaviour framework as currently described in [I-D.ietf-pcn-marking-behaviour] in order to enable the Single Threshold marking approach. Threshold marking in this draft conforms to [I-D.pcn-marking-behaviour] except for "intermediate depth". We propose an addition of two marking parameters: Marking.frequency, which is a ratio between actual marking and marking according to PCN-threshold marking.

5. Acknowledgements

This research was partially supported by the National Institute of Information and Communications Technology (NICT), Tokyo, Japan.

We thank Mr. Takayuki Uchida at U-software Corporation for his helping simulations and giving technical advices, and Professor Shigeo Shioda at Chiba University for his fruitful discussion.

6. Appendix A: Simulation Setup and Environment

6.1. Network and Signaling Models

We used three types of network topologies shown in the Figures below for simulations. They are the same as those in [I-D.charny-pcn-single-marking] and the first two figures are the same as those in [I-D.briscoe-.tsvwg-cl-phb]. The first type of network topology is Single Link, as shown in Fig. A.1. The second type is a multi-link network with a single bottleneck (termed "RTT"), as shown in Fig.A.2. The third type is a range of multi-bottleneck topologies shown in Fig.A.3 (termed "Parking Lot").

A single link between an ingress and an egress node is shown in Fig.A.1, where all flows enter at node A and depart from node B. This topology is used for the basic verification of the behavior of the algorithms with respect to a single IEA in isolation.
As shown in Fig. A.2, a set of ingresses (A;B;C) are connected to an interior node in the network (D). This topology is used to study the behavior of the algorithm where many IEAs share a single bottleneck link. The number of ingresses varied in different simulation experiments in the range of 10-1000. All links generally have different propagation delays. So this propagation delays were chosen randomly in the range of 1ms - 100 ms (although in some experiments all propagation delays are set the same). This node D in turn is connected to the egress (F). In this topology, different sets of flows between each ingress and the egress converge on a single link D-F, where the PCN algorithm is enabled. The capacities of the ingress links are not limiting, and hence, no PCN is enabled on those. The bottleneck link D-F is modeled with a 10ms propagation delay in all simulations. Therefore, the range of round-trip delays in the experiments is from 40ms to 220ms.

Another type of network of interest is parking lot (PLT) topology, which has multi-bottleneck links. The simplest PLT with 2 bottlenecks is illustrated in Fig A.3(a). A traffic matrix with this network on this topology is as follows:

- an aggregate of "2-hop" flows entering the network at A and leaving at C (via the two links A-B-C)
o an aggregate of "1-hop" flows entering the network at D and leaving at E (via A-B)

o an aggregate of "1-hop" flows entering the network at E and leaving at F (via B-C).

In the 2-hop PLT shown in Fig. A.3 (a) the points of congestion are links A-B and B-C. The capacity of all other links is not-limited. We also experiment with larger PLT topologies with 3 bottlenecks (see Fig. A.3(b)) and 5 bottlenecks (Fig. A.3 (c)). In all cases, we simulated one ingress-egress pair that carries the aggregate of "long" flows traversing all N bottlenecks (where N is the number of bottleneck links in the PLT topology), and N ingress-egress pairs that carry flows traversing a single bottleneck link and exiting at the next "hop". In all cases, only the "horizontal" links in Fig. A.3 were the bottlenecks, with non-limiting capacities of all "vertical" links. Propagation delays for all links in all PLT topologies are set to 1 ms.

Our simulations concentrated primarily on the range of capacities of "bottleneck" links with sufficient aggregation - above 128 Mbps for voice and 2.4Gbps Mbps for SVD. It should generally be expected that the higher link speeds will result in higher levels of aggregation, and hence generally better performance of the measurement-based algorithms. Therefore it seems reasonable to believe that the studied link speeds do provide meaningful evaluation targets.

In the simulation model, a call request arrives at the ingress and immediately sends a message to the egress. The message arrives at the egress after the propagation time plus link processing time (but no queuing delay). When the egress receives this message, it immediately responds to the ingress with the current CLE. If the CLE is below the specified CLE-threshold, the call is admitted, otherwise it is rejected. An admitted call sends packets according to one of the chosen traffic models for the duration of the call (see next section). Propagation delay from source to the ingress and from destination to the egress is assumed to be negligible and is not modeled.

Admissible and Supportable rates are half link speed and 90% of link speed, respectively, in all the links. Actual queue size is 80,000 packets, which is not the byte size because queue size is set by only packet in NS2 simulator.

6.2. Traffic Models

We use the same types of traffic as those of [I-D. briscoe-tsvwg-cl-phb], [I-D. charny-pcn-single-marking]. These are CBR voice, on-off
traffic approximating voice with silence compression (we termed "VBR"), and on-off traffic with higher peak and mean rates (we termed the latter "Synthetic Video" (SVD). On-off traffic (VBR and SVD) is described with two state Markov chain and on and off periods were exponentially distributed with the specified mean.

Each flow arrives according to the Poisson process and the distribution of flow duration was chosen to be exponentially distributed with a mean of 1 min, regardless of the traffic type, which is the same as that in [I-D. briscoe-tsvwg-cl-phb] and [I-D. charny-pcn-single-marking].

Traffic parameters for each type are summarized below:

CBR voice
- Packet length: 160 bytes
- Packet inter-arrival time: 20ms \((160*8)/(64*1000)\) sec
- Average rate: 64Kbps

On-off traffic approximating voice with silence compression
- Packet length: 160 bytes
- Packet inter-arrival time during On period: 20ms
- Long-term average rate: 21.76 Kbps
- On period mean duration: 340ms; during the on period traffic is sent with CBR voice parameters described above
- Off period mean duration 660ms; no traffic is sent for the duration of the off period

SVD
- Packet length: 1500 bytes
- Packet inter-arrival time during On period: 1ms
- Long term average rate: 4 Mbps
- On period mean duration: 340 ms
- Off period mean duration: 660 ms
6.3. Performance Metrics

In all our experiments, we use the percent deviation of the mean rate achieved in the experiment from the expected load level as performance metric. We term these "over-admission" and "over-termination" percentages, depending on the type of the experiment.

In our experiments of admission control, we compare the actual achieved average throughput to the desired traffic load (admissible rate) and measure the standard deviation of throughput at 500 ms intervals to the desired traffic load. The desired traffic load is not admissible rate exactly because a signal packet for a new flow is sent for admission. Therefore, the desired traffic load is the amount of the signal packets including not admitted flows less than admissible rate.

6.4. Parameter setting for STM

All the simulations were run with the following values:

- TB1.size = 20 ms at Admissible rate,
- TB1.threshold = 10ms at Admissible rate,
- TB2.size = 10ms at Supportable rate,
- TB2.threshold = 8ms at Supportable rate,
- Marking.frequency = 1/3 and 1
- CLE threshold = 0.05 (less than Marking.frequency times ratio between admissible and supportable rates)
- CLEweight = 0.0005
- The number of sequential marked packets for termination = 100
- Extra percentage of receiving rate more than the difference between sending and receiving rates for termination at one time = 5.0
- Interval for measuring sending rate or receiving rate = 100ms.

6.5. Parameter setting for CL

All the simulations were run with 5ms at Supportable rate as token bucket size for Termination and the following Virtual Queue thresholds:
6.6. Simulation Environment

We used NS2 for our simulation experiments. Simulations were run on Red Hat Enterprise Linux (64bit), Dell Power Edge 1950, Intel Quad-core Xeon 2.66GHz, 32GB RAM.

7. Appendix B: Admission Control Simulation

7.1. Parameter setting for admission control

We evaluate over-admission percentages when the load of the bottleneck is five times of the Admissible rate, that is, 2.5 times of the link speed. The simulation time is 300 sec and simulation results during time interval between 200 sec and 300sec are used for performance metric to obtain the results in the steady state.

Egress measurement parameters for STM: In our simulations, the CLE is computed as an exponential weighted moving average (EWMA) with a weight of 0.0005. The CLE is computed on a per-packet basis.

Egress measurement parameters for CL: In our simulations, the CLEthreshold was chosen as 0.05. The CLE is computed as an exponential weighted moving average (EWMA) with a weight of 0.3. The CLE is computed on a per-packet basis. These values are the same as those in Sect. 8.2.3 in [I-D.charny-pcn-single-marking].

7.2. Basic evaluation

Over-admission percentage statistics are evaluated using the Single Link topology. Table B.1 gives results of an admission control simulation when the load is the five times the Admissible rate. When the traffic type is CBR, the link speed is 128Mbps and the load is the rate of 5000 flows. The call interval per Ingress-Egress Aggregate(IEA) is 0.012sec. When the traffic type is VBR, the link speed is 78.3Mbps and the load is the rate of 9000 flows. The call interval per Ingress-Egress Aggregate(IEA) is 0.0067sec. When the traffic type is SVD, the link speed is 2.45Gbps and the load is the...
rate of 1500 flows. The call interval per Ingress-Egress Aggregate (IEA) is 0.04sec. We show the average of results of five simulations with different random seeds each traffic type. The performance of STM is very good although it is inferior to that of CL. We repeated this simulation by using different random seeds.

<table>
<thead>
<tr>
<th>Type</th>
<th>Over Admission %</th>
<th>Standard deviation %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STM</td>
<td>CL</td>
</tr>
<tr>
<td>CBR</td>
<td>0.285</td>
<td>0.028</td>
</tr>
<tr>
<td>VBR</td>
<td>1.017</td>
<td>0.979</td>
</tr>
<tr>
<td>SVD</td>
<td>4.308</td>
<td>4.476</td>
</tr>
</tbody>
</table>

Table B.1: Over admission percentage and standard deviation statistics obtained with Single Link

7.3. Effect of Ingress-Egress Aggregation

We evaluated the effect of Ingress-Egress Aggregation using RTT topology. As simulations in [I-D. charny-pcn-single-marking], the aggregate load on the bottleneck is the same across each traffic type with the aggregate load being evenly divided among all ingresses.

7.3.1. CBR

Table B.2 shows simulation results with traffic load condition when the traffic type is CBR. In this case, link speed of the bottleneck is 128Mbps for all the case of number of ingresses.
<table>
<thead>
<tr>
<th>No. of ingresses</th>
<th>No. of connections per IE pair</th>
<th>Call Interval per IEA (sec)</th>
<th>Over Admission (%)</th>
<th>Standard deviation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>STM</td>
<td>CL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STM</td>
<td>CL</td>
</tr>
<tr>
<td>10</td>
<td>500</td>
<td>0.12</td>
<td>0.585</td>
<td>0.761</td>
</tr>
<tr>
<td>70</td>
<td>71</td>
<td>0.84</td>
<td>-1.656</td>
<td>0.727</td>
</tr>
<tr>
<td>600</td>
<td>8</td>
<td>7.20</td>
<td>-5.218</td>
<td>0.038</td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
<td>12.0</td>
<td>-6.030</td>
<td>-0.179</td>
</tr>
</tbody>
</table>

Table B.2: Over admission percentage statistics with CBR, RTT

7.3.2. VBR

Table B.3 shows simulation results with traffic load condition when the traffic type is VBR. In this case, link speed of the bottleneck is 78Mbps for all the case of number of ingresses.

<table>
<thead>
<tr>
<th>No. of ingresses</th>
<th>No. of connections per IE pair</th>
<th>Call Interval per IEA (sec)</th>
<th>Over Admission (%)</th>
<th>Standard deviation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>STM</td>
<td>CL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>STM</td>
<td>CL</td>
</tr>
<tr>
<td>10</td>
<td>900</td>
<td>0.07</td>
<td>-0.753</td>
<td>1.345</td>
</tr>
<tr>
<td>70</td>
<td>128</td>
<td>0.47</td>
<td>-3.249</td>
<td>1.462</td>
</tr>
<tr>
<td>600</td>
<td>15</td>
<td>4.00</td>
<td>-5.427</td>
<td>0.699</td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
<td>12.0</td>
<td>-5.870</td>
<td>-0.077</td>
</tr>
</tbody>
</table>

Table B.3: Over admission percentage statistics with VBR, RTT

7.3.3. SVD

Table B.4 shows simulation results with traffic load condition when the traffic type is SVD. In this case, link speed of the bottleneck is 2448Mbps for all the case of number of ingresses.
### Table B.4: Over admission percentage statistics with SVD, RTT

<table>
<thead>
<tr>
<th>No. of ingress</th>
<th>No. of connections per IE pair</th>
<th>Call Interval (sec)</th>
<th>Over Admission (%)</th>
<th>Standard deviation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>150</td>
<td>0.4</td>
<td>1.593</td>
<td>4.870</td>
</tr>
<tr>
<td>70</td>
<td>42</td>
<td>1.4</td>
<td>-1.981</td>
<td>4.898</td>
</tr>
<tr>
<td>600</td>
<td>10</td>
<td>5.6</td>
<td>-5.502</td>
<td>2.597</td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
<td>12.0</td>
<td>-7.192</td>
<td>0.221</td>
</tr>
</tbody>
</table>

7.4. Effect of multi-bottleneck

We evaluated the effect of multi-bottleneck with PLT topology. Over-admission percentage of each bottleneck ID is shown in Table B.5. Standard deviation of over admission with multibottleneck of each bottleneck ID is shown Table B.6. When CBR was used as traffic type, link speed of all the bottleneck is 128Mbps and call-interval per IEA is 0.024sec. When VBR was used as traffic type, link speed of all the bottleneck is 78Mbps and call-interval per IEA is 0.01sec. When SVD was used as traffic type, link speed of all the bottleneck is 2448Mbps and call-interval per IEA is 0.08sec.

<table>
<thead>
<tr>
<th>Traffic Type</th>
<th>Algorithm</th>
<th>Bottleneck LinkId 1</th>
<th>Bottleneck LinkId 2</th>
<th>Bottleneck LinkId 3</th>
<th>Bottleneck LinkId 4</th>
<th>Bottleneck LinkId 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBR</td>
<td>STM</td>
<td>1.587</td>
<td>1.589</td>
<td>1.617</td>
<td>1.565</td>
<td>1.694</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>0.019</td>
<td>0.025</td>
<td>0.015</td>
<td>0.005</td>
<td>0.009</td>
</tr>
<tr>
<td>VBR</td>
<td>STM</td>
<td>2.567</td>
<td>2.320</td>
<td>2.168</td>
<td>2.269</td>
<td>2.341</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>0.560</td>
<td>0.540</td>
<td>0.585</td>
<td>0.614</td>
<td>0.620</td>
</tr>
<tr>
<td>SVD</td>
<td>STM</td>
<td>2.541</td>
<td>2.279</td>
<td>2.537</td>
<td>2.297</td>
<td>2.594</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>1.887</td>
<td>1.908</td>
<td>1.952</td>
<td>2.123</td>
<td>2.382</td>
</tr>
</tbody>
</table>

Table B.5: Over admission percentage with multibottleneck
<table>
<thead>
<tr>
<th>Traffic Type</th>
<th>Algorithm</th>
<th>Bottleneck LinkId</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBR</td>
<td>STM</td>
<td>2.270</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>1.002</td>
</tr>
<tr>
<td>VBR</td>
<td>STM</td>
<td>5.293</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>2.385</td>
</tr>
<tr>
<td>SVD</td>
<td>STM</td>
<td>6.474</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>6.042</td>
</tr>
</tbody>
</table>

Table B.6: Standard deviation of over admission with multibottleneck

7.5. Fairness among different Ingress-Egress pairs

In [I-D.charny-pcn-single-marking], the fairness is illustrated using the ratio between bandwidth of the long-haul aggregates and the short-haul aggregates. We used CBR traffic with the load of each bottleneck is five times of the Admissible rate. Table B.7 summarizes the fairness results at different topologies. We measure ratios of average throughput of short-haul aggregates to that of long-haul aggregates during the simulation time at the first bottleneck. We display the average of the five simulations. The same link speed and call interval conditions in Table B.5 are used in this experiment.

<table>
<thead>
<tr>
<th>Topo</th>
<th>Simulation Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td>STM</td>
<td>PLT2</td>
</tr>
<tr>
<td>CL</td>
<td></td>
</tr>
<tr>
<td>STM</td>
<td>PLT3</td>
</tr>
<tr>
<td>CL</td>
<td></td>
</tr>
<tr>
<td>STM</td>
<td>PLT5</td>
</tr>
<tr>
<td>CL</td>
<td></td>
</tr>
</tbody>
</table>

Table B.7: Fairness performance
8. Appendix C: Flow termination

We evaluate over-termination percentages when the load of the bottleneck is 1.0, 1.5 and 3.0 times of link speed. The load was lower than the Admissible rate at the beginning of the simulation. The simulation time is 100 sec. The load was changed at half the simulation time, that is 50 sec. Simulation results during time interval between 70 sec and 100 sec are used for performance metric to obtain the results in the steady state. The rate of one IE-pair was changed from 0 to the load of 1.0, 1.5 and 3.0 times of link speed at bottleneck links with no admission control within 2.0 sec. This simulated changing of a route when there was a failure. When a failure happened, a new Ingress-Egress pair was generated. Their traffic was regarded as traffic added by changing of a route. The flows of the new Ingress-Egress pair were generated according to the Poisson process whose rate was the value of the generated flows divided by 2 sec. The flow generation stopped at the generated flows whose bandwidth is the same as the load. The duration time of all the flows is infinity.

8.1. Basic evaluation

Over termination percentages are evaluated with the Single Link topology. Table C.1 gives results of a termination control simulation when the load is the 1.0, 1.5 and 3.0 times Link speed. When the traffic type is CBR, the link speed is 40.5 Mbps and the load is the rate of 633, 950, and 1900 flows. When the traffic type is VBR, the link speed is 41.1 Mbps and the load is the rate of 1886, 2830, and 5660 flows. When the traffic type is SVD, the link speed is 613 Mbps and the load is the rate of 153, 230, and 460 flows. We show the average of results of five simulations with different random seeds each traffic type.
<table>
<thead>
<tr>
<th>Traffic Type</th>
<th>Load (x Link speed)</th>
<th>Over Termination %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>STM</td>
</tr>
<tr>
<td>CBR</td>
<td></td>
<td>5.54</td>
</tr>
<tr>
<td>VBR</td>
<td>1.0</td>
<td>17.66</td>
</tr>
<tr>
<td>SVD</td>
<td></td>
<td>17.59</td>
</tr>
<tr>
<td>CBR</td>
<td></td>
<td>4.60</td>
</tr>
<tr>
<td>VBR</td>
<td>1.5</td>
<td>28.34</td>
</tr>
<tr>
<td>SVD</td>
<td></td>
<td>33.43</td>
</tr>
<tr>
<td>CBR</td>
<td></td>
<td>3.86</td>
</tr>
<tr>
<td>VBR</td>
<td>3.0</td>
<td>30.82</td>
</tr>
<tr>
<td>SVD</td>
<td></td>
<td>38.21</td>
</tr>
</tbody>
</table>

Table C.1: Over termination percentage statistics obtained with Single Link

9. **Appendix D: Another Flow Termination Control**

When PCN-supportable rate is much less than the Physical rate, the flow termination control described in Sect.?? takes much time to control. We describe another flow termination control. A PCN-ingress node measures sending rate per PCN-egress nodes all the time. If a PCN-egress node receives L-sequential marked packets, it sends a message of termination to the PCN-ingress node. When the PCN-ingress node receives the message, it terminates flows according to x-% of sending rate. PCN-ingress node ignore the message during a time interval after termination. The value of x is small and the time interval is short.

10. **IANA Considerations**

    TBD

11. **Security Considerations**

    TBD
12. Informative References

[I-D.babiarz-pcn-3sm]

[I-D.briscoe-tsvwg-cl-phb]

[I-D.charny-pcn-single-marking]

[I-D.ietf-pcn-architecture]

[I-D.ietf-pcn-baseline-encoding]

[I-D.ietf-pcn-marking-behaviour]

[I-D.westberg-pcn-load-control]

Authors’ Addresses

Daisuke Satoh
NTT Advanced Technology Corporation
1-19-18, Nakacho
Musashino-shi, Tokyo  180-0006
Japan

Email: daisuke.satoh@ntt-at.co.jp

Mika Ishizuka
NTT Advanced Technology Corporation

Email: mika.ishizuka@ntt-at.co.jp
Oratai Phanachet
NTT Advanced Technology Corporation

Email: oratai.phanachet@ntt-at.co.jp

Yukari Maeda
NTT Advanced Technology Corporation

Email: yukari.maeda@ntt-at.co.jp
Full Copyright Statement

Copyright (C) The IETF Trust (2008).

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.