PCN Encodings

Steps towards a final decision?

Toby Moncaster
Georgios Karagiannis
Michael Menth
Bob Briscoe
Motivation

- Aim of this talk is to discuss open issues raised in previous talk
- Next stage aims to meet charter milestone for standards track encoding document
- Aim of this presentation is to move forward on making decision
- draft-chan will lead to informational RFC tracking history of decision making
Encoding Requirements

- **CL**
  - Unmarked, admission-marked, termination-marked
- **SM**
  - Unmarked, termination-marked
- **3sm**
  - No-pre-congestion, admission-stop, excess-traffic
- **LC-PCN**
  - Unmarked, Affected_Marked, PCN_Marked

Thus maximum of 3 encoding states required (A, B, C) requiring the following transitions: A → B → C → A

PCN encodings
Terminology

• There are a wide range of terminologies in use for PCN. This presentation will use the following (not necessarily my preferred option):

**PCN**  traffic is in PCN traffic class (charter requires this to be indicated using a DSCP)

**NP**  Not Precongested – traffic in PCN class that hasn’t been pre-congestion marked

**AM**  Admission Marked – indicate ingress to stop admission

**TM**  Termination Marked – indicate need to terminate flows

**Af.M**  Affected Marked – indicate traffic that shares path with marked traffic
3 classes of encodings

All valid encoding options belong to one of 3 classes:
1. Use only DSCP codepoints
2. Use one DSCP codepoint + other codepoints from ECN
3. Use two DSCP codepoints + limited ECN codepoints

• Each of these will be described briefly
• Then will introduce major constraints
• Then discuss pros and cons
• Then attempt to reach consensus on which option the WG favours
1) Using only DSCPs

- Each of the states will be given a different DSCP.
- All DSCPs indicate traffic is PCN
- Use 2 or 3 DSCPs to indicate the PCN states (as required by the particular solution)

2) Using 1 DSCP & ECN

- DSCP indicates traffic is PCN. ECN codepoints indicate which PCN state
- Numerous variations proposed.
- draft-chan-pcn-encoding-comparison-03 lists several variants
2) Using 1 DSCP & ECN

- DSCP indicates traffic is PCN. ECN codepoints indicate which PCN state
- Numerous variations proposed. Following is an example
3) Using 2 DSCPS & limited ECN

- DSCP indicates traffic is PCN. Limited ECN codepoints in conjunction with DSCP indicate which PCN state
- This proposal very new and not yet discussed in draft-chan-pcn-encoding-comparison-03
- Following slides introduce motivation behind this particular encoding class.
Tunneling – A **MAJOR** constraint

<table>
<thead>
<tr>
<th>incoming inner</th>
<th>incoming outer</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>00 10 01 11</td>
</tr>
<tr>
<td>00</td>
<td>00 00 00 drop</td>
</tr>
<tr>
<td>10</td>
<td>10 10 10 11</td>
</tr>
<tr>
<td>01</td>
<td>01 01 01 11</td>
</tr>
<tr>
<td>11</td>
<td>11 11 11 11</td>
</tr>
</tbody>
</table>

- This constrains us not to use 00, 01 or 10 for carrying AM or TM (as these will get lost on decapsulation)
- Also can’t use 00 for NP since if the inner header is 00 & outer is 11 decapsulator drops the packet
### 3) Using 2 DSCPs & limited ECN

- **CE-bits in inner header of a tunnelled packet cannot be overwritten by the decaps node**
  - Transition from 11 to 01/10 cannot be preserved by decaps node
  - Only the following transitions possible

- **Codepoint for NP MUST reach codepoints for AM and TM**
  - Reuse of 01/10

- **Codepoint for AM should minimize impact of DSCP switching on ECMP**
  - Use the same DSCP
  - Reuse of 11
  - ⇒ No rerouting for PCN rate < PCN upper threshold (normal operation)

- **Only 1 reachable codepoint for TM**

---

<table>
<thead>
<tr>
<th>DSCP1</th>
<th>DSCP2</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Transition Diagram" /></td>
<td><img src="image2.png" alt="Transition Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DSCP1</th>
<th>01/10*</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>AM</td>
<td></td>
</tr>
<tr>
<td>TM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DSCP2</th>
<th>01/10*</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NP</td>
<td>AM</td>
</tr>
<tr>
<td></td>
<td>TM</td>
<td></td>
</tr>
</tbody>
</table>
Pros & cons for option 1

All DSCPs

Pros:
- Support for tunnels
- Allows transparent carrying of ECN

Cons:
- Requires at least 2 DSCPs (for 2 codepoint solutions) require at least 2n DSCPs for n precedence classes
- Possibility of undesirable interactions with ECMP
- Requires router to check existing marking before applying new marking (mustn’t remark TM → AM) for some solutions
Pros & cons for option 2

1 DSCP and ECN field

Pros:
- Plentiful codepoints

Cons:
- Impossible with the tunneling constraint identified above
Pros & cons for option 3

2 DSCPs + partial ECN

Pros:
- Support for some tunnels
- Requires less DSCPs requires 2n DSCPs for n precedence classes
- Not susceptible to ECMP interactions under normal operation

Cons:
- ECN can only be carried using IP in IP tunnels or moving it to a non-PCN class
- Possible ECMP interactions with TM marking
Pros and Cons - Discussion

• Previous slides not exhaustive list of pros/cons
• Have decided to ignore issue of leakage as only relevant to misconfigured routers
• Option 2 *could* become feasible if we re-write rules on encapsulation of ECN – see Bob’s ID (currently on hold):
  draft-briscoe-tsvwg-ecn-tunnel-00.txt
  – However this is too long term for initial PCN charter…
• Need contributions from floor about which encoding choice is best

DISCUSS!