TRILL ECN Support

draft-eastlake-trill-ecn-support-01.txt

Donald E. Eastlake, 3rd
d3e3e3@gmail.com

Bob Briscoe
ietf@bobbriscoe.net
**ECN Background**

- ECN propagates
  - “ECN-capable transport” (ECT) down
  - Congestion Experienced (CE) up
- ECT is necessary for incremental deployment
  - See IP ECN codepoints table (right)

- Similar incremental deployment problem for TRILL
  - if legacy *egress* does not understand ECN
  - will not propagate upward to forwarded IP inner header
  - would black-hole congestion signals

### IP-ECN codepoint table (right)

<table>
<thead>
<tr>
<th>IP-ECN codepoint</th>
<th>value</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not-ECT</td>
<td>00</td>
<td>Not ECN-capable transport</td>
</tr>
<tr>
<td>ECT(0)</td>
<td>10</td>
<td>ECN-Capable Transport</td>
</tr>
<tr>
<td>ECT(1)</td>
<td>01</td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>11</td>
<td>Congestion Experienced ('marked')</td>
</tr>
</tbody>
</table>

---

**July 2016**

**TRILL ECN Support**
Adding ECN Support to TRILL Header
Extension Flags Word

<table>
<thead>
<tr>
<th>TRILL-ECN codepoint</th>
<th>value</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not-ECT</td>
<td>00</td>
<td>Not ECN-capable transport</td>
</tr>
<tr>
<td>ECT(0)</td>
<td>10</td>
<td>ECN-Capable Transport</td>
</tr>
<tr>
<td>ECT(1)</td>
<td>01</td>
<td></td>
</tr>
<tr>
<td>NCCE</td>
<td>11</td>
<td>Non-Critical Congestion Experienced</td>
</tr>
</tbody>
</table>

Non-Critical Hop-by-Hop Bits
Critical Ingress-to-Egress Bits

Critical Congestion Experienced CCE flag
Defer mark / drop decision to egress

Possible ECN marking

End Station

Ingress RBridge

Transit

Egress RBridge

End Station

Copy traffic ECN field to Non-Critical Hop-by-Hop TRILL ECN field (not necessary for ‘Classic’ [RFC3168] ECN, but needed for variants.)

“Tunnel”

Congested transit RBridge that supports ECN marks using CCE flag (Critical Ingress-to-Egress)

No checking for ECT

Possible ECN marking

No ingress or transit dependence on Egress capabilities

Two cases:
1. Egress supports ECN decap, correctly combines outer TRILL ECN with inner IP ECN [RFC6040]
2. Egress doesn't understand ECN, drops any frame with a CltE bit set (Default behavior, which is desired)
Recap: ECN tunnelling rules at egress  
[ RFC6040 ]

<table>
<thead>
<tr>
<th>incoming inner</th>
<th>Arriving TRILL 3-bit ECN codepoint</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not-ECT</td>
</tr>
<tr>
<td>Not-ECT</td>
<td>Not-ECT</td>
</tr>
<tr>
<td>ECT(0)</td>
<td>ECT(0)</td>
</tr>
<tr>
<td>ECT(1)</td>
<td>ECT(1)</td>
</tr>
<tr>
<td>CE</td>
<td>CE</td>
</tr>
</tbody>
</table>

Outgoing header

TRILL egress same as  
[ RFC6040 ] but 3 TRILL ECN bits. So map 3 bits to the 4 codepoints as shown in table:

<table>
<thead>
<tr>
<th>NCHbH TRILL ECN</th>
<th>CItE CCE</th>
<th>Arriving TRILL 3-bit ECN</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>0</td>
<td>Not-ECT</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>ECT(0)</td>
</tr>
<tr>
<td>01</td>
<td>0</td>
<td>ECT(1)</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>CE</td>
</tr>
<tr>
<td>00</td>
<td>1</td>
<td>CE</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>CE</td>
</tr>
<tr>
<td>01</td>
<td>1</td>
<td>CE</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>CE</td>
</tr>
</tbody>
</table>
Changes in Draft -00 > -01

- Last IETF: presented 3 possible solutions
  - draft-00 wrote up solution #2
  - re-written to specify chosen solution: #3
- Renamed two CE fields to:
  - Non-Critical Congestion Experienced
  - Critical Congestion Experienced
- Added section on support for ECN variants
  - pre-congestion notification (PCN)
  - L4S (successful BoF on Tuesday)
Adding support to TRILL for Low Latency Low Loss Scalable throughput (L4S)

- for background on L4S see:
  - draft-briscoe-aqm-duaq-coupled, draft-briscoe-tsvwg-ecn-l4s-id

- On transit TRILL RBridge classify on TRILL-ECN field, then

  - Classic queue:
    ```
    if (p > max(random(), random()) ) {
      mark(frame, CCE )
    }
    ```

  - L4S queue:
    ```
    if (p > random() ) {
      if (p' > random() ) mark(frame, CCE )
      else mark(frame, NCCE )
    }
    ```

- then deferring mark/drop decision to egress gives desired outcome
- without any L4S logic at the egress
Next Steps

• Review the draft please
  • comprehensibility
  • Implementability

• WG adoption call
End

Donald E. Eastlake, 3\textsuperscript{rd}
d3e3e3@gmail.com

Bob Briscoe
ietf@bobbriscoe.net