PACKET PWE3 – EFFICIENT FOR IP/MPLS

IETF80 (PRAGUE) – MAR 27- APR 1, 2011
SRIGANESH KINI
DAVID SINICROPE
Problem statement

› Need for a packet service that can carry any protocol (similar to an Ethernet pseudowire)

› Service should be efficient for the most common protocol carried by the service

› IP and MPLS are pre-dominant protocols in many deployments

› Packet service efficient for IP/MPLS is useful
Service model

› Service is modeled as a VPWS that can carry packets of any protocol
  - Henceforth referred to as Packet Pseudowire (PPW)

› Efficient encapsulation defined for PPW over an MPLS PSN
  - Henceforth referred to as “Packet Pseudowire – Efficient for IP/MPLS” (PPW-EIM)
Network Reference model

› A single layer-2 (virtual) circuit is an access circuit (AC) to a PPW-EIM

› An AC of a PPW-EIM must not encapsulate another layer-2 circuit. E.g. in a Q-in-Q scenario, S-tag cannot be an AC to a PPW-EIM since it has multiple C-tags
Solution – with Control-Word

› CW is used to signal whether the packet is of type - IP, MPLS or ‘other’

› IP and MPLS packets encapsulated in PW without layer-2 header

› For non IP/MPLS packets, the layer-2 header is included. The protocol type in the layer-2 header indicates the layer-3 protocol type.
Solution – without CW

› Packet following bottom of label-stack is always IP

› Since there is a single bottom-of-stack bit in MPLS label stack, MPLS packets don’t need special identification

› For non IP/MPLS packets, an IP header encap (GRE) is used for the entire packet (including layer-2 header). A non-routable IP address is used as destination IP address to indicate that packet is non IP/MPLS.

› Even if intermediate nodes hash based on IP header there is no re-ordering.
Example: Router interconnect

- R1, R2 – routers running LLDP, ISIS on the inter-connecting p2p IP/MPLS interface
- PE1, PE2 – PEs providing VLL service using PPW-EIM
- IP/MPLS traffic encapsulation in MPLS PSN has no layer-2 header
- PE1 and PE2 encapsulate LLDP, ISIS packets (including layer-2 header) into GRE (if no CW is used) or following the CW (when CW is used)
Conclusion

› Lesser bandwidth used.

› Fragmentation is reduced for jumbo IP/MPLS packets

› Multi-layer network inefficiency reduced

› Enables flow based applications to parse packets efficiently even if there are multiple layers.
  - This includes ability to do ECMP based on IP (a widely deployed capability today)
v/s draft-ietf-pwe3-packet-pw (virtual ethernet)

› Advantages
  – Lesser bytes on the wire (Bandwidth efficient)
  – Less chance of fragmentation (throughput efficient)
  – IP ECMP is possible (even for multi-layer networks)
  – FAT-PW is not necessary for ECMP

› Disadvantages
  – Not possible to carry layer-2 circuit encapsulated inside a layer-2 circuit
  – For the no CW case – GRE encapsulation is more involved
Draft positioning

- This draft is positioned as an enhancement to using the ethernet PW as a packet-PW.
### Efficiency analysis

<table>
<thead>
<tr>
<th>IP Payload size (end-user data)</th>
<th>Total packet size in PSN</th>
<th>Efficiency as %age gain in end-PPW-EIM (no CW) v/s PWE3-ETH - PWE3-ETH (with CW)+flo</th>
<th>Efficiency as %age of bandwidth PPW-EIM (no CW) v/s PWE3-ETH - PPW-EIM (with CW)+flo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PWE3-ETH</td>
<td>PPW-EIM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without CW</td>
<td>with CW</td>
</tr>
<tr>
<td><strong>IP Payload size</strong></td>
<td></td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td><strong>AC Payload Size - IP</strong></td>
<td></td>
<td>60</td>
<td>96</td>
</tr>
<tr>
<td><strong>Ethernet Header size</strong></td>
<td></td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td><strong>802.1q single tag</strong></td>
<td></td>
<td>100</td>
<td>104</td>
</tr>
<tr>
<td><strong>802.1ad double tag</strong></td>
<td></td>
<td>104</td>
<td>108</td>
</tr>
<tr>
<td><strong>IP Payload size</strong></td>
<td></td>
<td>128</td>
<td>148</td>
</tr>
<tr>
<td><strong>AC Payload Size - IP</strong></td>
<td></td>
<td>148</td>
<td>184</td>
</tr>
<tr>
<td><strong>Ethernet Header size</strong></td>
<td></td>
<td>184</td>
<td>188</td>
</tr>
<tr>
<td><strong>802.1q single tag</strong></td>
<td></td>
<td>188</td>
<td>192</td>
</tr>
<tr>
<td><strong>802.1ad double tag</strong></td>
<td></td>
<td>192</td>
<td>196</td>
</tr>
<tr>
<td><strong>IP Payload size</strong></td>
<td></td>
<td>512</td>
<td>532</td>
</tr>
<tr>
<td><strong>AC Payload Size - IP</strong></td>
<td></td>
<td>532</td>
<td>568</td>
</tr>
<tr>
<td><strong>Ethernet Header size</strong></td>
<td></td>
<td>568</td>
<td>572</td>
</tr>
<tr>
<td><strong>802.1q single tag</strong></td>
<td></td>
<td>572</td>
<td>576</td>
</tr>
<tr>
<td><strong>802.1ad double tag</strong></td>
<td></td>
<td>576</td>
<td>580</td>
</tr>
</tbody>
</table>
Comments
Welcome