ECMP in Existing Bridges

May 2012

Ben Mack–Crane
Requests for ECMP in Existing Bridges

Network operators wanting ECMP also want:
  - Controlled upgrade process (F-TAG support)
  - To use existing infrastructure where possible

This raises the question:
  - Can existing bridges support ECMP?
Existing Bridge Functions

- Existing ports do not support F-TAG
  - No Flow Hash
  - No TTL

- Existing relay entity does not support hash-based lookup (multiple ports in Dynamic Filtering Entry)

- How can existing bridges support ECMP?
Region Requirements

- Consistent VLAN control mode assignment (STP, PBB-TE, SPBM, SPBV)
- Consistent ECT-ALGORITHM assignment for SPB VLANs
- If existing bridges are to support ECMP they must support SPBM & ECMP ECT Algorithm.
ECMP Support in Existing Relay Entity

- FDB selection based on DA + VID
- For each individual address with multiple next hop choices, use hash to select one to install in FDB (Dynamic Filtering Entry)

- “Per address” ECMP granularity.

<table>
<thead>
<tr>
<th>Address</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
</tr>
</tbody>
</table>
ECMP in an Existing Network

- Address ECMP can work in existing network (no change in port or relay capabilities)
- Loop prevention by Agreement Protocol (or infrastructure topology)
Migration to Flow–based ECMP

- F–TAG capability must exist at edge nodes
- Core nodes may still use address ECMP
  - No F–TAG processing
Migration to Flow–based ECMP

- Core node upgrade to support F–TAG can be incremental
  - Consistent F–TAG processing within each Bridge

Flow Hash Distribution

Equal Cost Hops to A,B,C,D
Can leave responsibility with operator, or

Can provide **automatic mode selection**

- Flow-based mode selected iff all edge nodes support F-TAG (select for each ECMP Base VID)
Automatic Mode Selection

- Base VID X selects flow-based ECMP (F-TAG)
- Base VID Y selects address ECMP (no F-TAG)
- Core nodes use address ECMP
Automatic Mode Selection Benefits

- Prevents F-TAG leaking out of SPT Region
  - Could cause hard-to-diagnose problems
- Automatically selects best mode
- Simplifies upgrade process
Automatic Mode Selection Support

- Advertise F-TAG capability
  - Node level (SPB Instance sub-TLV)
  - Port level (ISID-ADDR sub-TLV)
- ISIS-SPB calculation includes mode selection

<table>
<thead>
<tr>
<th>Type (3)</th>
<th>Octet</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Length</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>B-MAC Address</td>
<td>3-8</td>
<td>6</td>
</tr>
<tr>
<td>reserved</td>
<td>9</td>
<td>4 bits</td>
</tr>
<tr>
<td>Base VID</td>
<td>9-10</td>
<td>12 bits</td>
</tr>
</tbody>
</table>

Use a bit to indicate F-TAG capability at this address (port)

<table>
<thead>
<tr>
<th>Type (1)</th>
<th>Octet</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Length</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>CIST Root Identifier</td>
<td>3-10</td>
<td>8</td>
</tr>
<tr>
<td>CIST External Root Path Cost</td>
<td>11-14</td>
<td>4</td>
</tr>
<tr>
<td>Bridge Priority</td>
<td>15-16</td>
<td>2</td>
</tr>
<tr>
<td>reserved</td>
<td>17-18</td>
<td>11 bits</td>
</tr>
<tr>
<td>V</td>
<td>18</td>
<td>1 bit</td>
</tr>
<tr>
<td>SPSourceID</td>
<td>19-20</td>
<td>20 bits</td>
</tr>
<tr>
<td>Number of Trees</td>
<td>21</td>
<td>1</td>
</tr>
</tbody>
</table>

Use a bit to indicate F-TAG capability at this node
Proposal for ECMP in Existing Bridges

We are getting requests for:
- Controlled upgrade process (F-TAG support)
- To use existing infrastructure where possible

Supporting these requests seems possible:
- Some ECMP support possible in existing Bridges (address mode)
- Automatic mode selection can provide additional benefits
- Consider adding these to 802.1Qbp