Headroom Measurement
Protocol Design

Lily Lv
To-Do List

• Ethertype for Qdt
  ➢ What EtherType should be used for the round trip delay?

• DCBX: PFC Configuration TLV format design
  ➢ Important not to let packet formats and perceived encoding efficiencies to drive the protocol design.

• Managed objects
  ➢ The effort, implementation cost, and purpose of statistic gathering and retention requires careful consideration

• Timestamp point clarification
  ➢ Will (t3-t2) be impacted (variably) by queue delay?
  ➢ further specify t1, t4

• Timestamp accuracy
  ➢ What is the accuracy of t1, t4?
Ethertype for Qdt

Reuse Qcz (CI) Ethertype 89-A2

**Table 47-1—Layer-2 CIM Encapsulation**

<table>
<thead>
<tr>
<th>Octet</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDU EtherType (89-A2)</td>
<td>1</td>
</tr>
<tr>
<td>Version</td>
<td>3</td>
</tr>
<tr>
<td>Subtype</td>
<td>3</td>
</tr>
<tr>
<td>CIM PDU</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 47-4—CIM PDU**

<table>
<thead>
<tr>
<th>Octet</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>1</td>
</tr>
<tr>
<td>Reserved</td>
<td>1</td>
</tr>
<tr>
<td>Add/Del</td>
<td>1</td>
</tr>
<tr>
<td>destination_address</td>
<td>2</td>
</tr>
<tr>
<td>source_address</td>
<td>8</td>
</tr>
<tr>
<td>vlan_identifier</td>
<td>14</td>
</tr>
<tr>
<td>Encapsulated MSDU length</td>
<td>16</td>
</tr>
<tr>
<td>Encapsulated MSDU</td>
<td>18</td>
</tr>
</tbody>
</table>

**Subtype:**
This field, 4 bits in length, shall be transmitted with the value 0 to indicate an encapsulated CIM PDU. The Subtype field occupies the least significant 4 bits of the first octet of the layer-2 CIM Encapsulation.

**Qdt proposal**

<table>
<thead>
<tr>
<th>Octet</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDU Ethertype(89-A2)</td>
<td>1</td>
</tr>
<tr>
<td>Version</td>
<td>3</td>
</tr>
<tr>
<td>Subtype</td>
<td>3</td>
</tr>
<tr>
<td>Headroom Measurement PDU</td>
<td>4</td>
</tr>
</tbody>
</table>

**Subtype 0, CIM**

**Subtype 1, Headroom Measurement Message**

**Question:**
Is “65-529” too big for headroom measurement PDU?
PFC Configuration TLV format design (1/7)

• Protocol design for capability notification.
  • Augment DCBX by extending PFC configuration TLV

Phase 1: Capability notification
-- If both sides support PFC HDRM, initiate PFC HDRM Measurement Request, otherwise, stop the procedure.

Phase 1 answers:
- Which option(s) of measurement is supported?
- Which option of measurement will be used?
Should phase 1 answer
- What the compensation value is for each option?
PFC Configuration TLV format design (2/7)

- DCBX mechanism
  - DCBX has 3 types of attributes:
    - Informational attributes
    - Asymmetric attributes
    - Symmetric attributes
  - PFC configuration TLV is sent using symmetric attributes passing.
    - Symmetric attributes: “the passing of an attribute from one port to its peer port with objective of both ports utilizing the same attribute value.”
  - ‘Willing’ is important in symmetric attribute passing.
    - “A Willing port shall set its operational attribute to that indicated in the received TLV if the received TLV has the W bit set to zero. If both the local port and remote port are willing, then the attribute values of the port with the lower numerical MAC address shall take precedence.”

![Diagram of PFC configuration TLV format](image1)

![Symmetric state machine](image2)
PFC Configuration TLV format design (3/7)

- Non-PTP measurement required information from PFC configuration TLV
  - Non-PTP capability of remote port
  - (Compensation value is transmitted in response message)

![TLV format diagram]

Only capability is included in PFC configuration TLV. Compensation value will be included in request-response procedure.
PFC Configuration TLV format design (4/7)

- PTP-based measurement required information from PFC configuration TLV
  - PTP-based capability of remote
  - Compensation value (internal processing delay) of remote

PTP comp > 0, PTP-based measurement is capable, compensation value equals to ‘PTP comp’
PTP comp = 0, PTP-based measurement is incapable.
PTP-based measurement required information from PFC configuration TLV

Does it contradict the symmetric attribute passing rule?

Assuming below case.
Local para: \( W=0, \) PTP comp = ‘a’ (>0)
Remote para: \( W=1, \) PTP comp = ‘b’ (>0)

According to symmetric attribute passing rule, local operation parameter of PTP comp will still be ‘a’. However, what local system really needs to calculate headroom is ‘b’.

PTP-based capability and PTP comp cannot be combined together in PFC configuration TLV.
- PTP-based capability is symmetric attribute
- PTP comp is informational attribute
PFC Configuration TLV format design (6/7)

- Proposal:
  - PFC configuration TLV only includes ‘capability’

Question: if both non-PTP and PTP-based are capable, which will be used?

Option 1: Define priority of the 2 methods.

Option 2: Only one capability can be set in TLV.

Each bit indicates one capability.

Mode: 00 non-capable
      01 non-PTP
      10 PTP based
      11 reserved
PFC Configuration TLV format design (7/7)

Proposal:

- ‘PTP comp’ for PTP-based measurement passes to peer separately.

Option 1: Define a new informational TLV - **PFC informational TLV**

<table>
<thead>
<tr>
<th>TLV type</th>
<th>TLV information string length</th>
<th>802.1 OUI</th>
<th>802.1 subtype</th>
<th>PTP comp</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>12</td>
<td>00-80-C2</td>
<td>xx</td>
<td></td>
</tr>
</tbody>
</table>

7 bits 9 bits 3 octets 1 octet 8 octets

**TLV header** **TLV information string = 12 octets**

Option 2: Do not specify it in Qdt.

Mention the internal processing delay impact, but allow vendor specific way to implement.

DCBX informational attributes:

“Informational attributes are exchanged via LLDP without any participation in a DCBX state machine.”
Thanks