DCBX Framework

Joe Pelissier
az-pelissier-dcbx-framework-0409v2
DCBX Framework Overview

- The DCBX Framework provides three capabilities utilizing LLDP:
  - **Informational Passing:** the passing of information that is used in DCB applications or configuration
    - e.g., passing the number of traffic classes supported by a bridge to an attached end station may enable the end station to optimize its buffer resources
  - **Symmetric Parameter Passing:** the passing of a parameter from one port to its peer port with objective of both ports utilizing the same parameter
    - e.g., Priority Flow Control configuration
    - The “to” and “from” ports are administratively configured
      - e.g., passing configuration to and end station from a bridge
    - In addition, defense mechanisms are supported to ensure that mismatched configurations do not impact the fabric
      - Optional, required, or non-existent on a feature-by-feature basis
  - **Asymmetric Parameter Passing:** the passing of a parameter from one port to its peer port, or between both peer ports. In this case, the resulting configuration of each port may or may not match its peer
    - e.g., ETS bandwidth configuration
    - The directions in which the parameters are passed is administratively configured

- The Framework provides standard state machines for each of the parameter passing methods
Asymmetric Parameter Passing – Overview

- Two TLVs:
  
  **Configuration TLV**: Provides current operational state and Willing bit

  Willing bit is not used by state machine

  Provided to higher layers to provide an indication of the expected behavior of the remote port

  **Recommendation TLV**: Provides recommendation for the operational state of remote port

  Transmitted only if the local port is configured to make recommendations (in which case it is transmitted in all LLDP PDUs

  Transmitted regardless of the “willingness” of the remote port

  Recommendation Valid (RV) bit provides a mechanism for the local port to delay making recommendations until it receives additional information about the remote port (e.g. number of traffic classes supported and priority to priority group assignments).
Asymmetric Parameter Passing State Variables

**LocalWilling:** Indicates that the local port has been administratively configured to accept recommendations. This value is included in the Willing bit of Configuration TLVs transmitted by this port.

**OperParam:** The current operational state of the parameter on the local port. This value is included as the parameter in the Configuration TLVs.

**LocalAdminParam:** The administratively configured state for the parameter. This becomes the operational state of the parameter by default, and may be overridden if the local port accepts a recommendation from the remote port.

**RemoteParam:** The parameter received in the last Configuration TLV. This variable is set to NULL if the last LLDP PDU received did not contain the Configuration TLV or if no LLDP PDUs have been received.

**RV:** TRUE indicates that the RV (Recommendation Valid) bit was set in the last Recommendation TLV received. FALSE indicates that the RV bit was not set in the last Recommendation TLV received. NULL indicates either the TLV was not present in the last LLDP PDU or that no LLDP PDUs have been received.
## Asymmetric Parameter Passing – Example TLVs

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Info String Len=17</th>
<th>802.1 OUI 00-80-C2</th>
<th>802.1 Subtype = 9</th>
<th>Reserved</th>
<th>Willing</th>
<th>Priority Assignment Table</th>
<th>Priority Group Configured Bandwidth Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 bits</td>
<td>9 bits</td>
<td>3 octets</td>
<td>1 octet</td>
<td>7 bits</td>
<td>1 bit</td>
<td>4 Octets</td>
<td>8 Octets</td>
</tr>
</tbody>
</table>

### Configuration TLV

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Info String Len=13</th>
<th>802.1 OUI 00-80-C2</th>
<th>802.1 Subtype = 10</th>
<th>Reserved</th>
<th>RV</th>
<th>Recommended Priority Group Bandwidth Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 bits</td>
<td>9 bits</td>
<td>3 octets</td>
<td>1 octet</td>
<td>7 bits</td>
<td>1 bit</td>
<td>8 Octets</td>
</tr>
</tbody>
</table>

### Recommendation TLV
Asymmetric Parameter Passing State Machine

Begin

Init
OperParam = LocalAdminParam;
OperParam != LocalAdminParam && !(LocalWilling && RV == TRUE)
LocalWilling && RV == TRUE

RxRecommend
OperParam = RemoteParam;
OperParam != RemoteParam && LocalWilling && RV == TRUE
LocalWilling || RV != TRUE

Note: RV can take on three values: TRUE, FALSE, and NULL. NULL indicates that the Recommendation TLV was not part of the last LLDP PDU received, or that no LLDP PDUs have been received.
Asymmetric Parameter Passing Ladder Diagram

Port A

OperParm = LocalAdminParam, RV, RecoParam

OperParm = RecoParam, RV, RecoParam

OperParm = RecoParam, RV, RecoParam

Port B

Two transfers in one direction and one transfer in the other direction required to transfer the parameters
Symmetric Parameter Passing – Overview

- Single TLV used
  
  Always carries the current local operational state
  
  Carries a “Willing” (“W”) bit
  
  A port that sets the W is considered “Willing”
  
  A Willing port shall set its operational parameter to that indicated in the received TLV if the received TLV has the W bit set to zero
Symmetric Parameter Passing State Variables

**LocalWilling:** Indicates that the local port has been administratively configured to accept the parameter from the remote port. This value is included in the Willing bit of DCBX TLVs transmitted by this port.

**RemoteWilling:** TRUE indicates that the Willing bit was set in the last TLV received. FALSE indicates that the Willing bit was not set in the last TLV received. NULL indicates either the TLV was not present in the last LLDP PDU or that no LLDP PDUs have been received.

**OperParam:** The current operational state of the parameter on the local port. This value is included as the parameter in the DCBX TLV.

**RemoteParam:** Contains the value of the last parameter (i.e. the operational state of the remote port) received in the TLV. NULL indicates either the TLV was not present in the last LLDP PDU or that no LLDP PDUs have been received.

**LocalAdminParam:** The administratively configured state for the parameter. This becomes the operational state of the parameter by default, and may be overridden if the local port accepts the parameter from the remote port.
## Symmetric Parameter Passing – Example TLV

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Info</th>
<th>802.1 OUI</th>
<th>802.1 Subtype</th>
<th>W</th>
<th>Reserved</th>
<th>PFC Cap</th>
<th>PFC Enable</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>String Len=6</td>
<td>00-80-C2</td>
<td>Subtype = 11</td>
<td>1 octet</td>
<td>4 bits</td>
<td>3 bits</td>
<td>1 Octets</td>
</tr>
</tbody>
</table>

- **7 bits**
- **9 bits**
- **3 octets**
- **1 octet**
- **1 bit**
- **4 bits**
- **3 bits**
- **1 Octets**
Symmetric Parameter Passing State Machine

Begin

Init

OperParam = LocalAdminParam;

OperParam != LocalAdminParam && !(LocalWilling && RemoteWilling == FALSE)

LocalWilling && RemoteWilling == FALSE

RxRecommend

OperParam = RemoteParam;

OperParam != RemoteParam && LocalWilling && RemoteWilling == FALSE

LocalWilling || RemoteWilling != FALSE
Symmetric Parameter Passing Ladder Diagram

Port A

LocalAdminParam = x, Willing

LocalAdminParam = y, !Willing

LocalAdminParam = y, Willing

Port B
Symmetric Parameter Passing – Defense Option - Overview

- May apply to symmetric parameters that may be turned “on” or “off”
  
  For example, Congestion Notification
  
  Does not apply to all such parameters
  
  Reference the individual feature descriptions to determine if applicable

- Defense mechanism ensures that received frames do not harm the network if the remote end does not have the feature enabled
  
  For example, remap frames to different priorities, discard the frame, etc.
  
  The specific mechanism is defined on a feature-by-feature basis

- If implemented, a RDY bit is provided in the TLV for each feature that may be individually enabled / disabled
  
  For example, CN provides one bit for each priority

- A common state machine is provided for enabling / disabling the defenses

- A port may elect to delay transmission of frames until it detects that the remote port is RDY
**Symmetric Parameter Passing – Defense Option – State Variables and Functions**

**DefenseEnabled**: TRUE indicates that the local port has been administratively configured to enforce defenses. FALSE indicates that the local port has not been administratively configured to enforce defenses. For some features, defense is not optional. In these cases, DefenseEnabled is always TRUE.

**OperParam**: The current operational state of the parameter on the local port. This value is included as the parameter in the DCBX TLV. TRUE indicates that the feature is enabled. FALSE indicates that it is not.

**RemoteParam**: Contains the value of the last parameter (i.e. the operational state of the remote port) received in the TLV. TRUE indicates that the feature is enabled. FALSE indicates that it is not. NULL indicates either the TLV was not present in the last LLDP PDU or that no LLDP PDUs have been received.

**SetDefenses()**: SetDefenses(On) enables defenses for this port. SetDefenses(Off) disables defenses for this port.
## Symmetric Parameter Passing – Defense Option - Example TLV

<table>
<thead>
<tr>
<th>TLV Type =127</th>
<th>TLV Info String Len=7</th>
<th>802.1 OUI 00-80-C2</th>
<th>802.1 Subtype = 13</th>
<th>W</th>
<th>Reserved</th>
<th>CN Mask</th>
<th>RDY Mask</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 bits</td>
<td>9 bits</td>
<td>3 octets</td>
<td>1 octet</td>
<td>1 bit</td>
<td>7 bits</td>
<td>1 octet</td>
<td>1 octet</td>
</tr>
</tbody>
</table>
Symmetric Parameter Passing – Defense Option - State Machine

Begin

<table>
<thead>
<tr>
<th>DefensesOn</th>
<th>DefensesOff</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDY=0;</td>
<td>RDY=1;</td>
</tr>
<tr>
<td>SetDefenses(On);</td>
<td>SetDefenses(Off)</td>
</tr>
<tr>
<td>!DefenseEnabled</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DefenseEnabled & OperParam & RemoteParam != TRUE
Symmetric Parameter Passing – Defense Option - Ladder Diagram

Port A  |  Port B
---|---
LocalAdminParam = FALSE, Willing, RDY
LocalAdminParam = TRUE, !Willing, !RDY
LocalAdminParam = TRUE, Willing, RDY
LocalAdminParam = TRUE, !Willing, RDY
TLV Summary
Traffic Classes Supported TLV

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Info String Len</th>
<th>802.1 OUI 00-80-C2</th>
<th>802.1 Subtype = 8</th>
<th>Reserved</th>
<th>Traffic Classes Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 bits</td>
<td>9 bits</td>
<td>3 octets</td>
<td>1 octet</td>
<td>5 bits</td>
<td>3 bits</td>
</tr>
</tbody>
</table>

- Generally provided as information
- E.g. may be used to:
  - Optimize buffer allocation (especially in a NIC)
- Informational Passing
ETS Configuration TLV

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Info String Len=17</th>
<th>802.1 OUI 00-80-C2</th>
<th>802.1 Subtype = 9</th>
<th>Reserved</th>
<th>Willing</th>
<th>Priority Assignment Table</th>
<th>Priority Group Configured Bandwidth Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 bits</td>
<td>9 bits</td>
<td>3 octets</td>
<td>1 octet</td>
<td>7 bits</td>
<td>1 bit</td>
<td>4 Octets</td>
<td>8 Octets</td>
</tr>
</tbody>
</table>

- Generally provided as information
- E.g. may be used to:
  - Generate a recommendation for PG Bandwidths
- Operates as the Configuration TLV with Asymmetric Parameter Passing
  - The Bandwidth table may be configured using Asymmetric Parameter Passing
- Priority Assignment table has one 4-bit entry per Priority
  - 0-7 indicate the Priority is mapped to the corresponding PG,
  - 13,14 indicate mapped to one of the AVB groups
  - 15 indicates priority operates as specified in 802.1Q (strict priority)
- Priority Group Configured Bandwidth Table has one 8-bit entry per PG
  - (always contains 8 entries).
  - Indicates the current BW configured for each Priority Group
  - Total must equal 100 (implies valid range is 0-100 for each entry)
ETS Recommendation TLV

- Provides a recommendation of how far end should configure the PG bandwidths
- This is the recommendation TLV used in Asymmetric Parameter Passing
- RV (Recommendation Valid), indicates that the TLV contains a recommendation. Set to zero if no recommendation is being provided at this time (although one may be provided in the future, e.g. after receipt of other TLVs)
- Priority Group Configured Bandwidth Table has one 8-bit entry per PG (always contains 8 entries) (reserved if RV set to 0).
  - Indicates the current BW configured for each Priority Group
  - Total must equal 100 (implies valid range is 0-100 for each entry)
    - If total does not equal 100, TLV is malformed and shall be ignored
  - If a device is “willing”, it may update its configured BWs in accordance with this TLV
    - OK for both ends to be “willing”

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Info String Len=13</th>
<th>802.1 OUI 00-80-C2</th>
<th>802.1 Subtype = 10</th>
<th>Reserved</th>
<th>RV</th>
<th>Recommended Priority Group Bandwidth Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 bits</td>
<td>9 bits</td>
<td>3 octets</td>
<td>1 octet</td>
<td>7 bits</td>
<td>1 bit</td>
<td>8 Octets</td>
</tr>
</tbody>
</table>

TLV Type =127
TLV Info String Len=13
802.1 OUI 00-80-C2
802.1 Subtype = 10
Reserved
RV
Recommended Priority Group Bandwidth Table
Priority-based flow Control TLV

<table>
<thead>
<tr>
<th>TLV Type =127</th>
<th>TLV Info String Len=6</th>
<th>802.1 OUI 00-80-C2</th>
<th>802.1 Subtype = 11</th>
<th>W</th>
<th>Reserved</th>
<th>PFC Cap</th>
<th>PFC Enable</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 bits</td>
<td>9 bits</td>
<td>3 octets</td>
<td>1 octet</td>
<td>1 bit</td>
<td>4 bits</td>
<td>3 bits</td>
<td>1 Octets</td>
</tr>
</tbody>
</table>

- Provides negotiation and information of PFC enabled / disabled per priority
- PFC Cap indicates the device’s limitation of how many traffic classes may simultaneously support PFC (not negotiated).
- Utilizes Symmetric Parameter Passing *without* the defense option
  - However, a defense option may yet be proposed
- PFC enable has 8 bits (one per priority)
  - A one indicates PFC is enabled on the priority
  - A zero indicates that PFC is disabled on the priority
  - Local policy in each end of the link decides whether to use the priority if the configuration does not match
Application Priority TLV

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Info String Len=variable</th>
<th>802.1 OUI 00-80-C2</th>
<th>802.1 Subtype = 12</th>
<th>W</th>
<th>Reserved</th>
<th>Application Priority Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 bits</td>
<td>9 bits</td>
<td>3 octets</td>
<td>1 octet</td>
<td>1 bit</td>
<td>7 bits</td>
<td>Multiple of 3 octets</td>
</tr>
</tbody>
</table>

- Provides indication of priority should be used for each protocol
- Uses Symmetric Parameter Passing *without* the defense option
- Application Priority Table contains zero or more entries:
  - 3 bits: Priority
  - 2 bits: Reserved
  - 3 bits: Sel (0: Ethertype, 1 Port# over TCP, 2 Port# over UDP, 3 Port# over TCP or UDP, 4 Port# over neither UDP nor TCP, 5-7 reserved for future use, devices shall ignore entries with these values)
  - 16 bits: Protocol ID (based on Sel field)
Congestion Notification TLV

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Info String Len</th>
<th>802.1 OUI</th>
<th>802.1 Subtype</th>
<th>W</th>
<th>Reserved</th>
<th>CN Mask</th>
<th>RDY Mask</th>
</tr>
</thead>
<tbody>
<tr>
<td>=127</td>
<td>=7</td>
<td>00-80-C2</td>
<td>=13</td>
<td>1</td>
<td>7 bits</td>
<td>1 octet</td>
<td>1 octet</td>
</tr>
</tbody>
</table>

- Provides synchronization and negotiation of Congestion Notification
- Uses the Symmetric Parameter Passing *with* defense mode *required*
  
i.e. the DefenseEnabled parameter in the Defense Option State Machine is always set TRUE
- CN Mask indicates if CN is enabled (1) / disabled (0) for corresponding priority
- RDY Mask indicates if defenses are enabled (0) or disabled (1) for a given priority. Reserved if corresponding CN Mask bit is 0
Thank You!