SSE-C3632S/SSE-C3632SR/SBM-25G-200

Data Centre Bridging Exchange

User's Guide

Revision 1.1

The information in this USER'S GUIDE has been carefully reviewed and is believed to be accurate. The vendor assumes no responsibility for any inaccuracies that may be contained in this document, makes no commitment to update or to keep current the information in this manual, or to notify any person organization of the updates. Please Note: For the most up-to-date version of this manual, please see our web site at www.supermicro.com.

`Super Micro Computer, Inc. (“Supermicro”) reserves the right to make changes to the product described in this manual at any time and without notice. This product, including software, if any, and documentation may not, in whole or in part, be copied, photocopied, reproduced, translated or reduced to any medium or machine without prior written consent.

DISCLAIMER OF WARRANTY ON SOFTWARE AND MATERIALS. You expressly acknowledge and agree that use of the Software and Materials is at your sole risk. FURTHERMORE, SUPER MICRO COMPUTER INC. DOES NOT WARRANT OR MAKE ANY REPRESENTATIONS REGARDING THE USE OR THE RESULTS OF THE USE OF THE SOFTWARE OR MATERIALS IN TERMS OF THEIR CORRECTNESS, ACCURACY, RELIABILITY, OR OTHERWISE. NO ORAL OR WRITTEN INFORMATION OR ADVICE GIVEN BY SUPER MICRO COMPUTER INC. OR SUPER MICRO COMPUTER INC. AUTHORIZED REPRESENTATIVE SHALL CREATE A WARRANTY OR IN ANY WAY INCREASE THE SCOPE OF THIS WARRANTY. SHOULD THE SOFTWARE AND/OR MATERIALS PROVE DEFECTIVE, YOU (AND NOT SUPER MICRO COMPUTER INC. OR A SUPER MICRO COMPUTER INC. AUTHORIZED REPRESENTATIVE) ASSUME THE ENTIRE COST OF ALL NECESSARY SERVICE, REPAIR, OR CORRECTION.

LIMITATION OF LIABILITY. UNDER NO CIRCUMSTANCES INCLUDING NEGLIGENCE, SHALL SUPER MICRO COMPUTER INC. BE LIABLE FOR ANY INCIDENTAL, SPECIAL, OR CONSEQUENTIAL DAMAGES THAT RESULT FROM THE USE OR INABILITY TO USE THE SOFTWARE OR MATERIALS, EVEN IF SUPER MICRO COMPUTER INC. OR A SUPER MICRO COMPUTER INC. AUTHORIZED REPRESENTATIVE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any disputes arising between manufacturer and customer shall be governed by the laws of Santa Clara County in the State of California, USA. The State of California, County of Santa Clara shall be the exclusive venue for the resolution of any such disputes. Super Micro's total liability for all claims will not exceed the price paid for the hardware product.

Manual Revision 1.1
Release Date: 08/11/2021

Document Revision History

Date	Revision	Description
12/08/2020	1.0	Initial document.
08/11/2021	1.1	Include SBM-25G-200

Contents

1 DCBX Overview

DCBX is a discovery & capability exchange protocol that is capable of discovering DCB compliant devices and exchange DCBX configuration information with them. Supermicro switches SSE-C3632S/SR and SBM-25G-200 support DCBX (Data centre Bridging Exchange) feature with version 1.01. The Converged Enhanced Ether (CEE) version of DCBX has a subtype of 2 and OUI of 0x001b21.

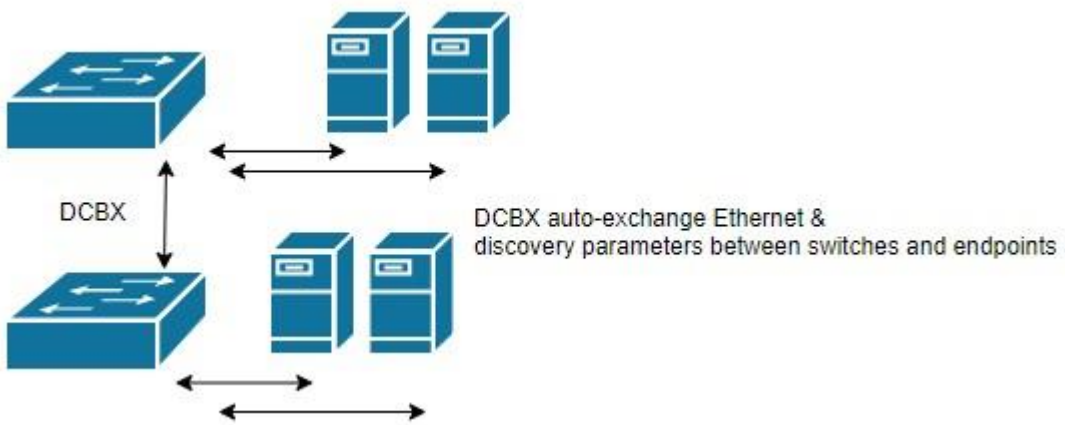
DCBX runs on the Physical Ethernet link between Supermicro switch (e.g SSE-C3632S) and Host Server's Network Card adapter (e.g. AOC-MH25G-m2S2T,AOC-S100G-m2C) that has DCBX capabilities.
DCBX protocol relies on the Link Layer Discovery Protocol (LLDP) to exchange DCBX information with its DCBX peer. DCBX peers (switch and the host adapter) negotiate the capabilities between them to send configuration values to the adapter. Auto PFC configuration from the switch to the host can be achieved with DCBX TLVs.

- Support DCBX version 1.01 (Converged Enhanced Ethernet version).
- DCBX capability on the switch remain disabled if the host network adapter does not support the DCBX.
- Please make sure that the DCBX feature is enabled on the host network adapter if it is not turned on by default.

2 DCBX Feature Benefits

Feature	Benefit Properties
Data Center Bridging Exchange (DCBX) Protocol	It allows exchange of Ethernet parameters between switches and Host Adapters.
Enhanced transmission selection	Provides bandwidth management between traffic types.
Priority Based Flow Control (PFC)	Provides a link level flow control mechanism that can be controlled independently for each frame priority.

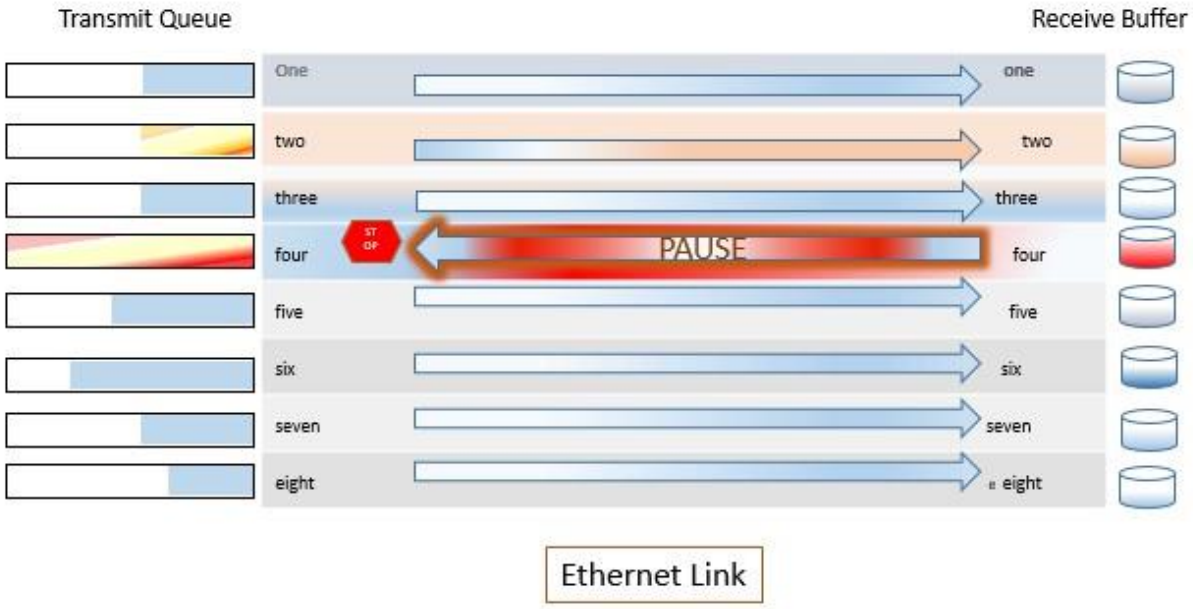
Data Center Bridging Exchange (DCBX) Protocol



Enhanced Transmission Selection (**ETS)** IEEE 802.1Qaz is a bandwidth management mechanism which enables us to allocate port bandwidth in such a way that maximizes bandwidth utilization for all flows on a link. ETS allows a port to share and re-allocate BW dynamically among its flows while guaranteeing a minimum amount of bandwidth to each flow.

Priority-based flow control (PFC), IEEE standard 802.1Qbb, is a link-level flow control mechanism. Which is an enhancement to the Ethernet pause mechanism, operates on single priority rather than pausing all traffic on a link.

PFC creates eight logically divided virtual links from A physical link and provides the capability to use pause on a single virtual link without affecting traffic on the other virtual links. PFC allows us to pause traffic selectively according to its class.



PFC Priority based Flow Control (Figure A)

3 DCBX Configuration Steps

Configuring DCBX involves the steps listed below.

1. Enable LLDP.
2. Create cee-map.
    a. Create a name for Priority (optional).
    b. Mark the application-protocol packets with required priority.
    c. Create a name for Priority-group (optional).
    d. Map priority to priority-group.
    e. Allocate bandwidth to the priority-group.
3. Apply cee-map to the interface.
4. Configure TLVs (optional).

Commands to configure the above steps on Super Micro Switch are given below in detail.

3.1 Enable LLDP feature on the switch

DCBX protocol relies on Link Layer Discovery Protocol (LLDP) to exchange information with peer. So LLDP must be enabled for DCBX feature to work.

Step	Command	Description
Step 1	**configure terminal**	Enters the configuration mode
Step 2	**Set lldp enable**	Enable and Configure LLDP
Step 3	**Exit**	Exit from configuration mode.

Example:

```
SMIS# configure terminal
SMIS(config)# set lldp enable
SMIS(config)# exit
```

3.2 Create cee-map

Converged Enhanced Ethernet map creates an association among application-protocol, priority, prioritygroup, and group-bandwidth. Four cee-maps can be created.

Step	Command	Description
Step 1	**configure terminal**	Enters the configuration mode
Step 2	**cee-map <CEE-map-id(1-4))>**	Creates a cee-map.
Step 3	**Exit**	Exit from configuration mode.

Example:

```
SMIS# configure terminal
SMIS(config)# cee-map 1
SMIS(config-cee-map)# exit
SMIS(config)# exit

SMIS# show cee-map 1
```

3.2.1 Create a description for Priority (optional)

There are 8 priorities available and they are identified by number 0 - 7. Creating a description for the priority helps to easily identify the traffic assigned to that priority. This step is optional and doesn't affect the functionality. Description has to be created before assigning the cee-map to the interface.

Step	Command	Description
Step 1	**configure terminal**	Enters the configuration mode
Step 2	**cee-map <CEE-map-id(1-4))>**	Creates a cee-map.
Step 3	**priority <pri(0-7)> description {<string(63)>}**	Creates a description for priority, which can be viewed in the show commands.
Step 4	**Exit**	Exit from configuration mode.

Example:

```
SMIS# configure terminal
SMIS(config)# cee-map 1
SMIS(config-cee-map)# priority 1 description "FTP traffic"
SMIS(config-cee-map)# exit
SMIS(config)# exit

SMIS# show cee-map 1
```

3.2.2 Create a name for Priority-group (optional)

CEE supports 9 priority-groups and they are identified by PGID number 0 – 7 and 15. Creating a description for the priority-group helps to easily identify the traffic assigned to that priority-group. This step is optional and doesn't affect the functionality. Description has to be created before assigning the cee-map to the interface.

Step	Command	Description
Step 1	**configure terminal**	Enters the configuration mode
Step 2	**cee-map <CEE-map-id(1-4))>**	Creates a cee-map.
Step 3	**group <id(0-7,15)> description {<string(63)>}**	Creates a description for priority group, which can be viewed in the show commands.
Step 4	**Exit**	Exit from configuration mode.

Example:

```
SMIS# configure terminal
SMIS(config)# cee-map 1
SMIS(config-cee-map)# group 0 description "Download Traffic"
SMIS(config-cee-map)# exit
SMIS(config)# exit

SMIS# show cee-map 1
```

3.2.3 Map priority to priority-group

Multiple priorities can be bundled together to form a priority-group. In other words, the traffic will be assigned to priority-group based on their priority. The priority-groups are identified by PGID number 0 – 7. There are 8 priorities available and they are identified by number 0 - 7. More than one priority can be assigned to a priority-group.

A default priority-to-priority-group mapping will be created when a cee-map is created. The default mapping is shown below.

Priority	Group	PFC	Description
0	0	No	LAN
1	0	No	
2	1	Yes	FCoE/FIP
3	0	No	
4	0	No	
5	0	No	
6	0	No	

7	0	No	

- Priority-group 15 is a special group; traffic shall be assigned to this group with no bandwidth limit, group-level PFC and members of this group are scheduled in strict priority order. It takes procedure over other the groups with non-zero bandwidth limit.

Step	Command	Description
Step 1	**configure terminal**	Enters the configuration mode
Step 2	**cee-map <CEE-map-id(1-4))>**	Creates a cee-map.
Step 3	**pri2pg <group(0-7, 15)> <> <> <> <> <> <> <>**	Maps the priority 0-7 respectively to the priority-groups.
Step 4	**Exit**	Exit from configuration mode.

Example:

```
SMIS# configure terminal
SMIS(config)# cee-map 1
SMIS(config-cee-map)# pri2pg 1 1 1 2 2 7 1 15
SMIS(config-cee-map)# end

SMIS# show cee-map 1
```

3.2.4 Allocate bandwidth to the priority-group

The 100% of bandwidth has to be divided as required and allocated among the 8 priority-groups. The total bandwidth allocated to 8 groups has to be 100.

Step	Command	Description
Step 1	**configure terminal**	Enters the configuration mode
Step 2	**cee-map <CEE-map-id (1-4))>**	Creates a cee-map.

Step 3	**group-bandwidth <bandwidth (0-100)> <> <> <> <> <> <> <>**	Allocates the bandwidth respectively to the priority-groups 0 - 7.
Step 4	**end**	Exit from configuration mode.

Example:

```
SMIS# configure terminal
SMIS(config)# cee-map 1
SMIS(config-cee-map)# group-bandwidth 10 10 10 20 10 10 10 20
SMIS(config-cee-map)# end
```

- Limited by ASIC scheduling hierarchy, the group bandwidth is implemented by per-priority basis. The port scheduling mode is Weighted Deficit Round Robin (WDRR). The priorities in the same group have the weight value by sharing bandwidth percentage. Each priority has weight 1 at least when the group doesn't allocate bandwidth.

- The bandwidth throughput depends on the weight rate of the total egress traffic.

Example:

In cee-map 1, priority 0 and 1 map to priority group 1. Priority 2, 3 and 4 map to group 2. Other priorities map to group 0. Priority group 1 has 30% bandwidth. Priority group 2 has 70% bandwidth. The weight is as below table

Priority	Group	Group Bandwidth percentage	Weight
0	1	30	30/2 = 15
1	1	30	30/2 = 15
2	2	70	70/3 + 1 = 24
3	2	70	70/3 = 23
4	2	70	70/3 = 23
5	0	0	1
6	0	0	1
7	0	0	1

Group 1 and 2 could have the bandwidth rate (30% and 70%) when

- The egress port transmits the traffic with the priority 0,1,2,3 and 4 only.
- Each incoming traffic load exceeds its egress bandwidth rate.

3.2.5 Enable PFC

Priority Flow Control (PFC) can be enabled for traffics based on priorities or priority-groups. PFC can be enabled/disabled in a cee-map before applying it to an interface.

Step	Command	Description
Step 1	**configure terminal**	Enters the configuration mode
Step 2	**cee-map <CEE-map-id(1-4))>**	Creates a cee-map.
Step 3	**PFC priority <pri(0-7)> {enable|disable}**	Enable/disable PFC for priority.
	PFC group <id(0-7)> {enable|disable}	Enable/disable PFC for priority-group.
Step 4	**end**	Exit from configuration mode.

- Default enable PFC for priority 3. Others are disabled.
- Default enable PFC for priority-group 1. Others are disabled.

Example:

```
SMIS# configure terminal
SMIS(config)# cee-map 1
SMIS(config-cee-map)# pfc priority 1 enable
SMIS(config-cee-map)# pfc group 2 enable
SMIS(config-cee-map)# end

SMIS# show cee-map 1

Use disable option to disable the PFC.
SMIS# configure terminal
SMIS(config)# cee-map 1
SMIS(config-cee-map)# pfc priority 1 disable
SMIS(config-cee-map)# pfc group 2 disable
```

```
SMIS(config-cee-map)# end
```

3.3 Apply cee-map to the interface

After the cee-map configuration is complete, it has to be applied to the physical interface for it to take effect. Same cee-map can be applied to multiple interfaces.

Step	Command	Description
Step 1	**configure terminal**	Enters the configuration mode
Step 2	**interface <interface-type> <interface-id>**	Enters the interface configuration mode. *Interface-type* –may be any of the following: fx-ethernet cx-ethernet
Step 3	**CEE <CEE-map-id(1-4))>**	Applies the cee-map.
Step 4	**dcbx cee**	Enables DCBX.
Step 5	**end**	Exit from configuration mode.

Example:

```
SMIS# configure terminal
SMIS(config)# interface fx-ethernet 0/1
SMIS(config-if)# cee 1
SMIS(config-if)# dcbx cee
SMIS(config-if)# end

SMIS# show interface fx-ethernet 0/1
```

3.4 Configure TLVs (optional)

Information such as DCBX control state, configuration, etc are exchanged between DCBX peers using Type Length Value (TLV) over LLDP protocol. Local operational configuration of each DCBX parameter is handled by DCBX state machine by comparing and synchronizing with the settings of its DCBX peer. TLV configuration are setup specific or NIC specific. The example show below are only for illustration purpose.

Apply CEE-map and enable DCBX before configuring TLVs.

Step	Command	Description
Step 1	**configure terminal**	Enters the configuration mode

Step 2	**interface <interface-type> <interface-id>**	Enters the interface configuration mode. *Interface-type* –may be any of the following: fx-ethernet cx-ethernet
Step 3	**CEE <CEE-map-id (1-4))>**	Applies the cee-map.
Step 4	**dcbx cee**	Enables DCBX.
Step 5	**LLDP TLV-select DCBX-CEE-PFC [advertise {on|off}] [willing {0|1}] [enable {0|1 }]**	Configures DCBX-CEE-PFC TLV.
Step 6	**LLDP TLV-select DCBX-CEE-pg [advertise {on|off}] [willing {0|1}] [enable {0|1}]**	Configures DCBX-CEE-pg TLV.
Step •	**end**	Exit from configuration mode.

Example:

```
SMIS# configure terminal
SMIS(config)# interface fx-ethernet 0/1
SMIS(config-if)# cee 1
SMIS(config-if)# dcbx cee
SMIS(config-if)# lldp tlv-select dcbx-cee-pfc advertise on willing 1
enable 0
SMIS(config-if)# lldp tlv-select dcbx-cee-pg  advertise on willing 1
enable 0
SMIS(config-if)# end

Use the no form of the command to remove the configuration; shown below
are some example.
SMIS# configure terminal
SMIS(config)# interface fx-ethernet 0/1
SMIS(config-if)# no lldp tlv-select dcbx-cee-pfc
SMIS(config-if)# no lldp tlv-select dcbx-cee-pg
SMIS(config-if)# no dcbx cee
SMIS(config-if)# no cee-map
```

4 Show commands for CEE-MAP and DCBX

Use ‘show cee-map’ to check the CEE-MAP configuration. This command displays the application-protocol to priority mapping, priority to priority-group mapping, and bandwidth allocation for the priority-groups.

show cee-map [<cee-map-id (1-4)>]

Example:

```
SMIS# show cee-map 1

CEE-Map 1

Ports : fx 0/1
Priority  Group  PFC  Description
-----------------------------------
0      1   No  LAN
1      1  Yes
2      1   No
3      2  Yes  FCoE/FIP
4      2   No
5      7   No
6      1   No
7      15   No

Group  Bandwidht(%)  PFC  Description
---------------------------------------
0            10   No  LAN
1            10  Yes  SAN
2            10   No
3            20   No
4            10   No
5            10   No
6            10   No
7            20   No
15          MAX   No

Application-Protocol-ID   Type       Protocol-ID    Priority
-------------------------------------------------------------
1   ether-type 0x8906         3
```

```
2    ether-type 0x8914          3
3    tcp-udp    3260            4
```

Use ‘show lldp dcbx’ command to check the current status/result of DCBX (CEE) use the below show command.

show lldp dcbx interface [<interface-type> <interface-id>]

Example:

```
SMIS# show lldp dcbx interface fx-ethernet 0/1
Fx0/1:
DCBX Control Message Exchange Information
-----------------------------------------
Status: Non-synchronized

Peer message seq#: 16777216 (acknowledged: 0)
Local message seq#: 2 (acknowledged: 16777216)

DCBX Feature Information
----------------------------------------
Feature: PG, Priority Groups
Type/subtype: 2/0
Enabled: Yes
Advertisement: Yes
Willing: No
Error: No
Operation status: Operational Config
(operation/desired/peer):
         PG0...10 / 10 / 10
         PG1...10 / 10 / 10
         PG2...10 / 10 / 10
         PG3...20 / 20 / 20
         PG4...10 / 10 / 10
         PG5...10 / 10 / 10
         PG6...10 / 10 / 10
         PG7...20 / 20 / 20
         PG15...MAX / MAX / MAX
         #TCs...8 / 8 / 8

Feature: PFC, Priority-based Flow Control
Type/subtype: 3/0
Enabled: Yes
```

```
Advertisement: Yes
Willing: No
Error: No
Operation status: Operational Config
(operation/desired.pg/peer):
         Pri0...1 / 0.1 / 1
         Pri1...0 / 0.0 / 0
         Pri2...0 / 0.0 / 0
         Pri3...1 / 1.0 / 1
         Pri4...0 / 0.0 / 0
         Pri5...0 / 0.0 / 0
         Pri6...0 / 0.0 / 0
         Pri7...1 / 0.1 / 1
         #TCs...8 / 8 / 8

Feature: Application Protocol
Type/subtype: 4/0
Enabled: Yes
Advertisement: Yes
Willing: No
Error: No
Operation status: Not Operational
(No support)
(operation/desired/peer):
         Operation Config
         Type       Protocol-ID    Priority
---------------------------------
ether-type 0x8906         3          ether-
type 0x8914        3         tcp-udp    3260
4

         Desired Config
         Type       Protocol-ID    Priority
---------------------------------
ether-type 0x8906         3          ether-
type 0x8914         3
         tcp-udp    3260
```

Other related show commands:

```
SMIS# show lldp neighbors
SMIS# show lldp neighbors detail
SMIS# show lldp traffic
SMIS# show lldp traffic [<iftype> <ifnum>]
```

5 Sample DCBX Configuration

The sample configuration shown below to configure priority group mapping/bandwidth and enable PFC for all priorities:

```
SMIS # show running-config
cee-map 3
 pri2pg 1 2 3 0 0 0 0 0
 pfc priority 0 enable
 pfc priority 1 enable
 pfc priority 2 enable
 pfc priority 3 enable
 pfc priority 4 enable
 pfc priority 5 enable
 pfc priority 6 enable
 pfc priority 7 enable
 group-bandwidth 0 40 60 0 0 0 0 0
 pfc group 0 enable
 pfc group 1 enable
 pfc group 2 enable
 pfc group 3 enable
 pfc group 4 enable
 pfc group 5 enable
 pfc group 6 enable
 pfc group 7 enable
 group 0 description Grp0000
 group 1 description Grp1111
 group 2 description Grp2222
 group 3 description Grp3333
 group 4 description Grp4444
 group 5 description Grp5555
 group 6 description Grp6666
 group 7 description Grp7777
!
interface cx-ethernet 0/3
 cee 1
 dcbx cee
```

```
!
interface fx-ethernet 0/7
 cee 1
 dcbx cee
!
SMIS# SMIS# show cee-map 1

CEE-Map 1

Ports :  cx-ethernet 0/3 fx-ethernet 0/7
Priority  Group  PFC  Description
-----------------------------------
       0      1  Yes  LAN
       1      2  Yes
       2      3  Yes
       3      0  Yes  FCoE/FIP
       4      0  Yes
       5      0  Yes
       6      0  Yes
       7      0  Yes

Group  Bandwidht(%)  PFC  Description
---------------------------------------
    0             0  Yes  Grp0000
    1            40  Yes  Grp1111
    2            60  Yes  Grp2222
    3             0  Yes  Grp3333
    4             0  Yes  Grp4444
    5             0  Yes  Grp5555
    6             0  Yes  Grp6666
    7             0  Yes  Grp7777
    15          MAX  Yes

Application-Protocol-ID   Type       Protocol-ID    Priority
-------------------------------------------------------------
                      1   ether-type 0x8906         3
                      2   ether-type 0x8914         3
                      3   tcp-udp    3260           4
SMIS#
```