L2 Openflow
ONF Update

Author: **Hesham ElBakoury**

Version: 1.0

HUAWEI TECHNOLOGIES CO., LTD.

HUAWEI

Agenda

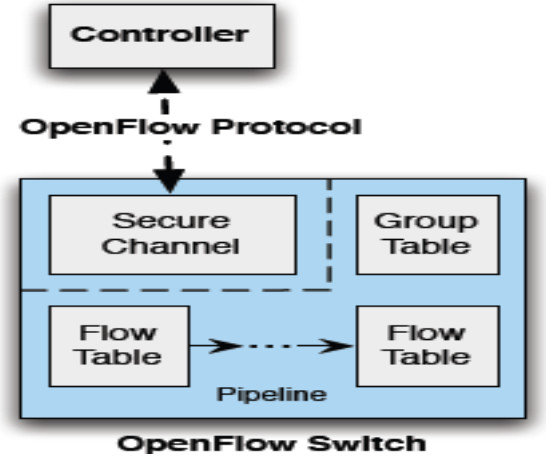
- **ONF Support for L2 Openflow**
- **Openflow Reliability Requirements**
  - Communication Reliability
  - Infrastructure Reliability
- **Using 1904.2 to address L2 Reliability Requirements**
  - LLC Encapsulation
- **L2 Shim Adaptation Layer**
- **ONF Open Source Projects.**
- **1904.2 Support for Infrastructure Reliability**
- **IEEE 802.1 Priority-Based flow Control**

ONF L2 Openflow Support

- **ONF decided to decouple Openflow protocol from the underlying reliable transport**
- **Chair of ONF Extensibility WG agreed to work with Huawei to determine the requirements of the underlying reliable transport protocol.**
  - Previous Huawei discussion with ONF on L2 openflow is recorded in https://rs.opennetworking.org/bugs/browse/EXT-463
  - The requirements will be specified in openflow 1.3.5 spec which is supposed to be released 1st half of 2015.

Openflow Reliability Requirements (1)

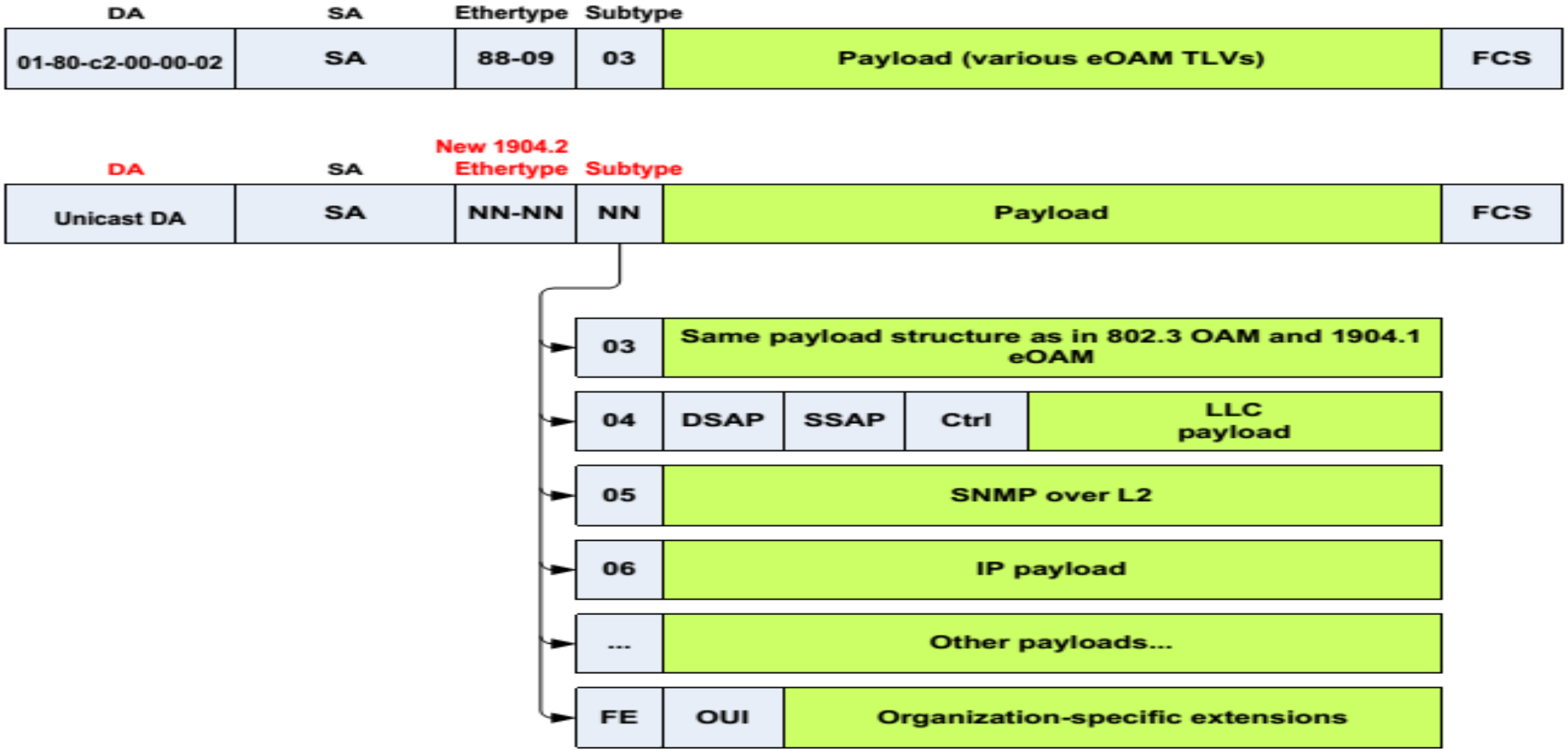
- The openflow controller system should not make the controller the single point of failure (addressed by Multi-controller support).
- Ensure the availability of control channels between the OF controller and OF switches and can continue to operate even during failure (can be addressed by 1904.2 for L2 Transport)
- The main control channel should provide reliable and secure communications with OF switches. The auxiliary connections could use unreliable transport like UDP.



Openflow Reliability Requirements (2)

- **For the main connection, if we do not use TCP/TLS, then the transport to be used must provide:**
  - Proper demultiplexing of multiple controller connections.
  - Retry mechanism. OpenFlow does not have any retry and assume the underlying mechanism is reliable.
  - Message ordering. The OpenFlow requires that message order be respected.
  - Flow Control

Current UMT Protocol Support



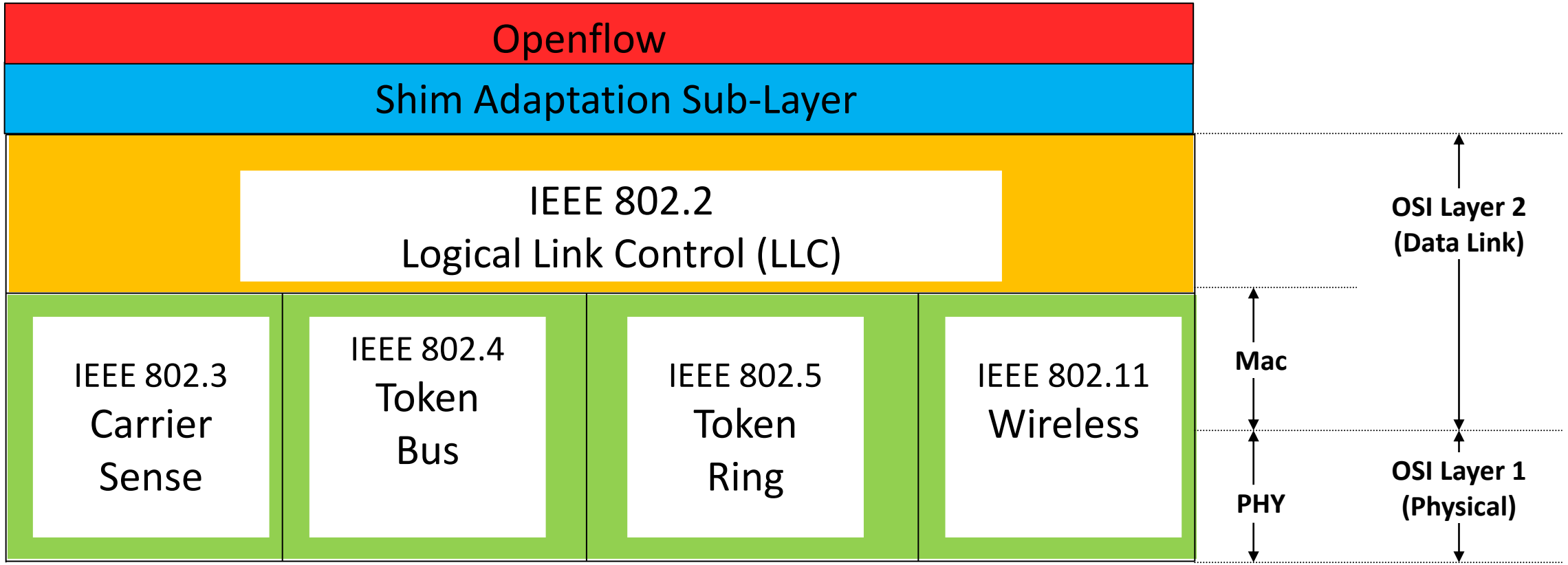
Openflow Packet Segmentation & Assembly

- How we handle PDU segmentation/assembly if intermediate switches between controller and OF switch do not support jumbo frames ?
  - OpenFlow requests don't consider the MTU and will routinely exceed the MTU size.
  - The chair of the extensibility group indicated that openflow should NOT do segmentation/assembly. This should be the function of the underlying transport layer.

Using LLC to Provide Reliable Communications

- **If LLC is used then ONF should request SSAP/DSAP from IEEE-RA.**
- **A shim adaptation sub-layer should be implemented between Openflow and LLC to perform the following functions:**
  - segmentation and assembly
  - convert northbound API with openflow to the southbound API with LLC.
  - Any other adaptation functions that may deem necessary

Using LLC Type 2 for Reliable L2 Communications



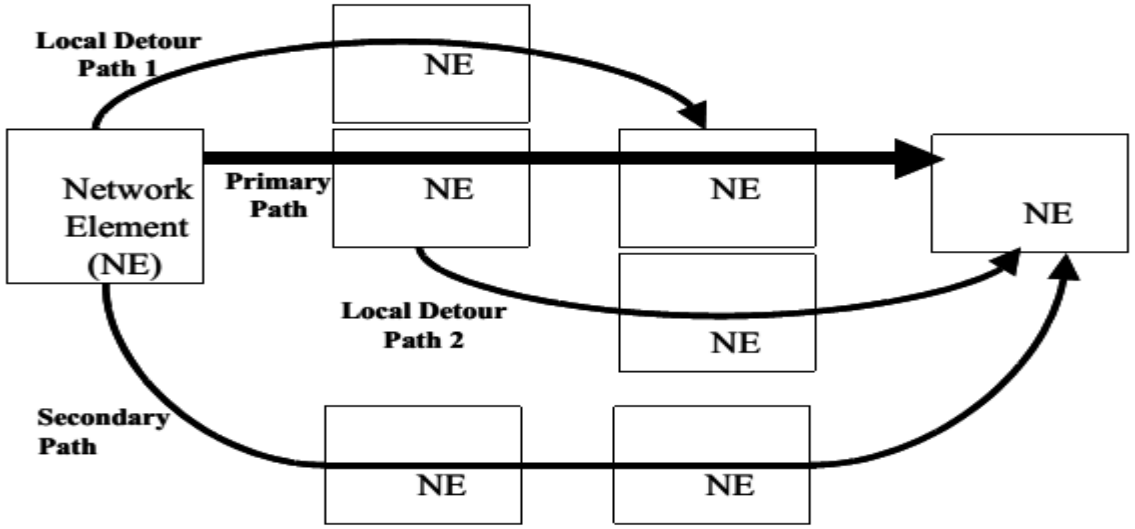
It has not been decided yet to specify in openflow spec a generic northbound API between Openflow and the underlying reliable transport

Using Open Source to Provide Reliable Communications

- **ONF encourage vendors to provide open source for different features & protocols.**
- **Open source project should be approved by ONF.**
- **There are potential two Open source projects:**
  - Develop code for reliable L2 communication protocol
    - **The new protocol should request an Ethertype from IEEE-RA or use UMT encapsulation**
  - Develop code for shim adaptation sub-layer if a protocol like LLC is used.
    - **The shim adaptation sub-layer should perform packet segmentation and assembly.**
    - **This sub-layer should support LLC API.**
- **Both projects should support northbound API with openflow.**

Infrastructure Protection Mechanisms

- Link Protection based on Link Aggregation
- PON Protection
- Aggregated Line and Node Protection (ALNP) service
- End-to-End Path Protection using Redundant Paths.
- Ethernet Protection can be achieved using mechanisms defined in G.8031/8032.
- IEEE 802.1ag/Y.1731 can be used for the etection of path failure/problems

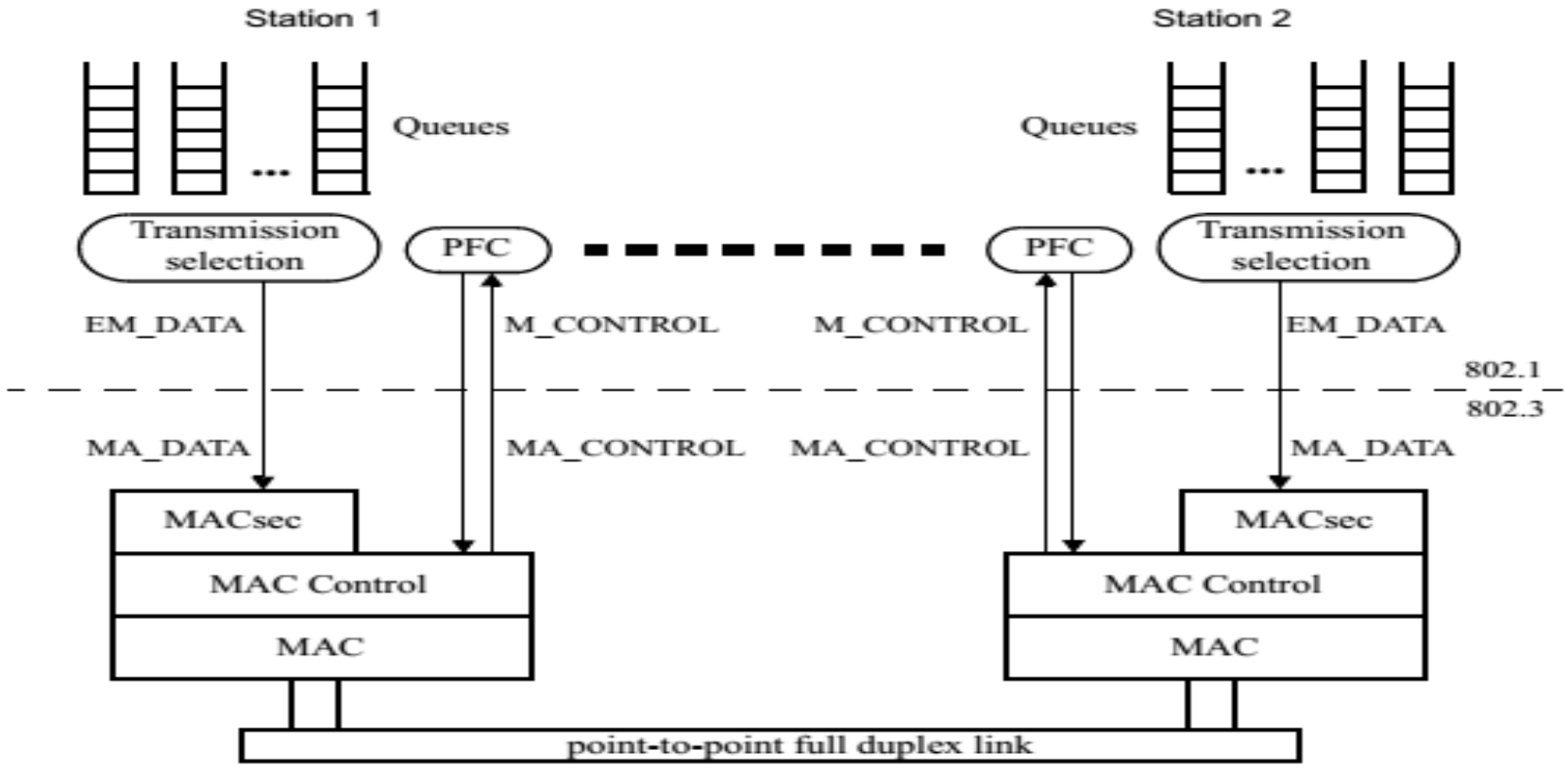


Priority-Based Flow Control

- Traditional IEEE 802.3 Ethernet defines an unreliable communication medium; it does not offer guarantees that a packet injected into the network will arrive at its intended destination.
- In a network path that normally consists of multiple hops between source and destination, lack of feedback between transmitters and receivers at each hop is one of the main causes of unreliability.
- For applications that cannot build reliability on upper layers, the addition of flow control functions at Layer 2 can offer a solution.
- IEEE 802.3x PAUSE suffers a basic disadvantage that limits its field of applicability: after a link is paused, a sender cannot generate any more packet.
- IEEE 802.1Qbb PFC extends the basic IEEE 802.3x PAUSE semantics to multiple CoSs, enabling applications that require flow control to coexist on the same wire with applications that perform better without it. PFCuses the IEEE 802.1p CoS values in the IEEE 802.1Q VLAN tag to differentiate up to eight CoS that can be subject to flow control independently

Priority-Based Flow Control (PFC) Peering

A system client wishing to inhibit transmission of data frames on certain priorities from the remote system on the link generates an
M_CONTROL.request primitive specifying:
a) The globally assigned 48-bit multicast address 01-80-C2-00-00-01;
b) The PFC opcode (i.e., 01-01); and
c) A request_operand_list with two operands indicating respectively the set of priorities addressed and the lengths of time for which it wishes to inhibit data frame transmission of the corresponding priorities.



PFC is a function defined only for a pair of full duplex MACs (e.g., 802.3 MACs operating in point-to-point full duplex mode) connected by one point-to-point link. Use of PFC on shared media such as EPON is out of the scope of this standard.

Thank you

www.huawei.com

Data Link Layer -- LLC

- **Logical Link Control (LLC)**
  - For IEEE 802, flow control, error control, and part of the framing duties are all brought together in this LLC sub-layer.
  - Framing is also performed in the MAC sub-layer.
  - The LLC provides a single data link control protocol for all IEEE LANs. (Contrast this with the MAC layer where we have different protocol versions for different LANs. See previous slide.)
  - The LLC makes the MAC sub-layer transparent.
  - Framing
    - **LLC provides a PDU similar to HDLC**

LLC Sublayer Connectionless Mode of Operation

- **The LLC sublayer standard, ISO/IEC 8802-2, describes two types of connection-less mode of operation for data communication between service access points:**
  - **Unacknowledged connectionless-mode (type 1),**
    - Information frames are exchanged between LLC entities without the need for the prior establishment of a logical link between peers. The LLC sublayer does not provide any acknowledgments for these LLC frames, nor does it provide any flow control or error recovery procedures.
    - LLC type 1 also provides a TEST function and an Exchange Identification (XID) function. The capability to act as responder for each of these functions is mandatory: This allows a station to check the functioning of the communication path between itself and any other station, to discover the existence of other stations, and to find out the LLC capabilities of other stations.
  - **Acknowledged connectionless-mode (type 3).**
    - With type 3 operation, information frames are exchanged between LLC entities without the need for the prior establishment of a logical link between peers. However, the frames are acknowledged to allow error recovery and proper ordering. Further, type 3 operation allows one station to poll another for data
    - This type has little or no commercial use.

LLC Sublayer Connection-mode of Operation

- **Connection-mode (type 2):**
  - A logical link is established between pairs of LLC entities prior to any exchange of information frames.
  - In the data transfer phase of operation, information frames are transmitted and delivered in sequence.
  - Error recovery using positive ACK and retransmission. F
  - low control using fixed-length sliding window of 8 frames are provided within the LLC sublayer.
- **Systems on the same IEEE 802.2 network may use 802.2 Type 2 communications after verifying that it is supported by both ends. This is accomplished using the 802.2 XID mechanism**

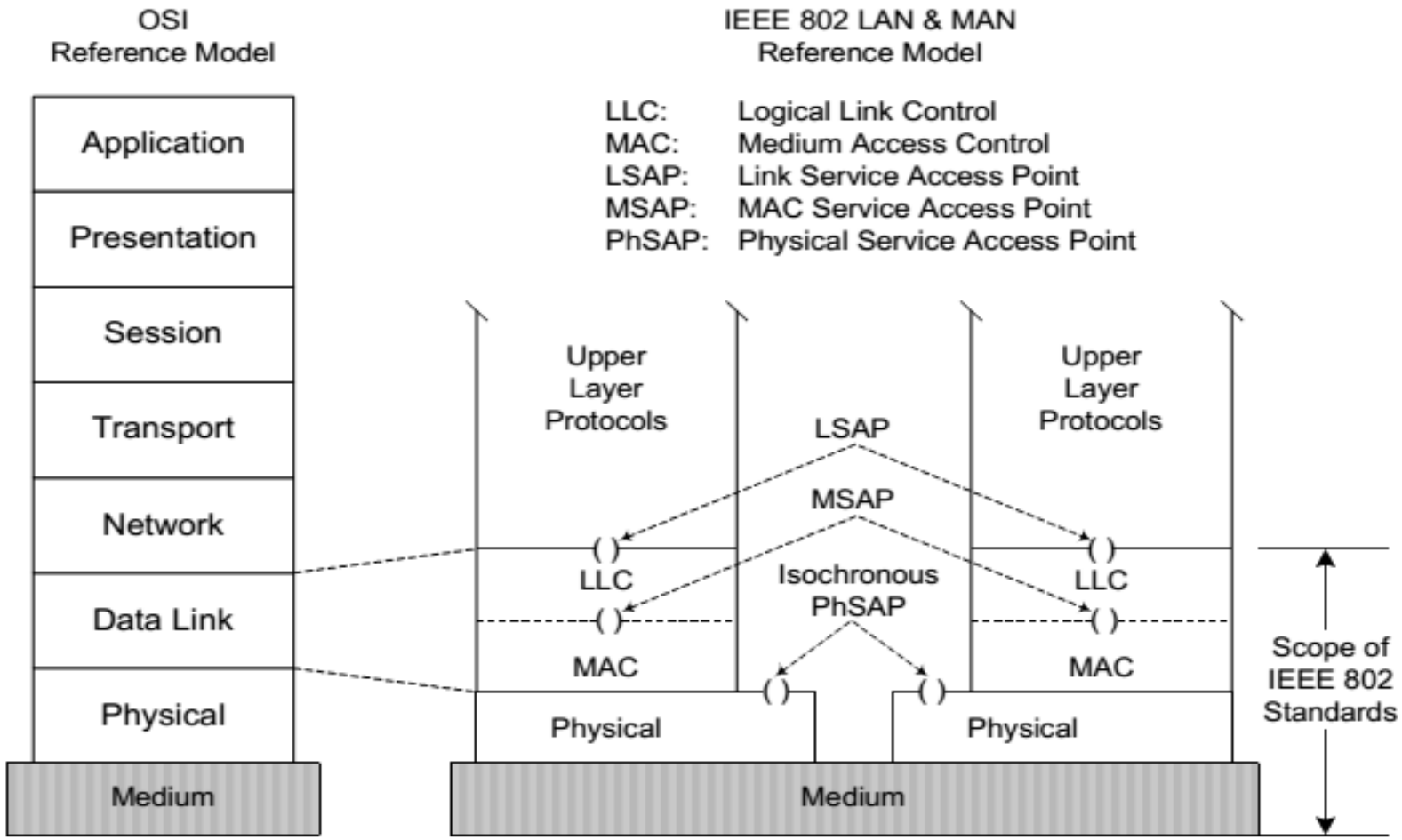
User Data Encapsulation

- **Type Field Encoding**
  - The length/type field (2 bytes) in the Ethernet frame contains a type value in the range of 0x600 to 0xFFFF indicating the higher-layer protocol that is being encapsulated.
  - The most common encapsulation used on Ethernet.
  - Ethertype values are registered with IEEE: http://standards.ieee.org/develop/regauth/ethertype/
  - Type field assignments are administered by the Registration Authority, IEEE Standards Department, P.O. Box 1331, 445 Hoes Lane, Piscataway, NJ 08855-1331, USA.
- **LLC Encoding**
  - The SSAP and DSAP identify the higher layer protocol.
  - The control field carries a value indicating the type of frame.
    - **LLC type1 uses unnumbered Information and provides connectionless, best effort service.**
    - **LLC type 2 provides reliable, connection-oriented services.**
  - SSAP/DSAP are allocated by IEEE-RA for a protocol which is a standard published by an internationally recognized standards organization, and has a potentially large field of application.

User Data Encapsulation (2)

- Clients using Length Encapsulation and LLC can communicate among themselves, and Clients using Type Encapsulation can communicate among themselves on the same LAN.
- When Type Encapsulation is used LLC Protocol is not used and does not need to be present.
- If a device supports some clients that use Type Encapsulation and others that use Length Encapsulation, the MAC can demultiplex frames to both sets of clients simultaneously.

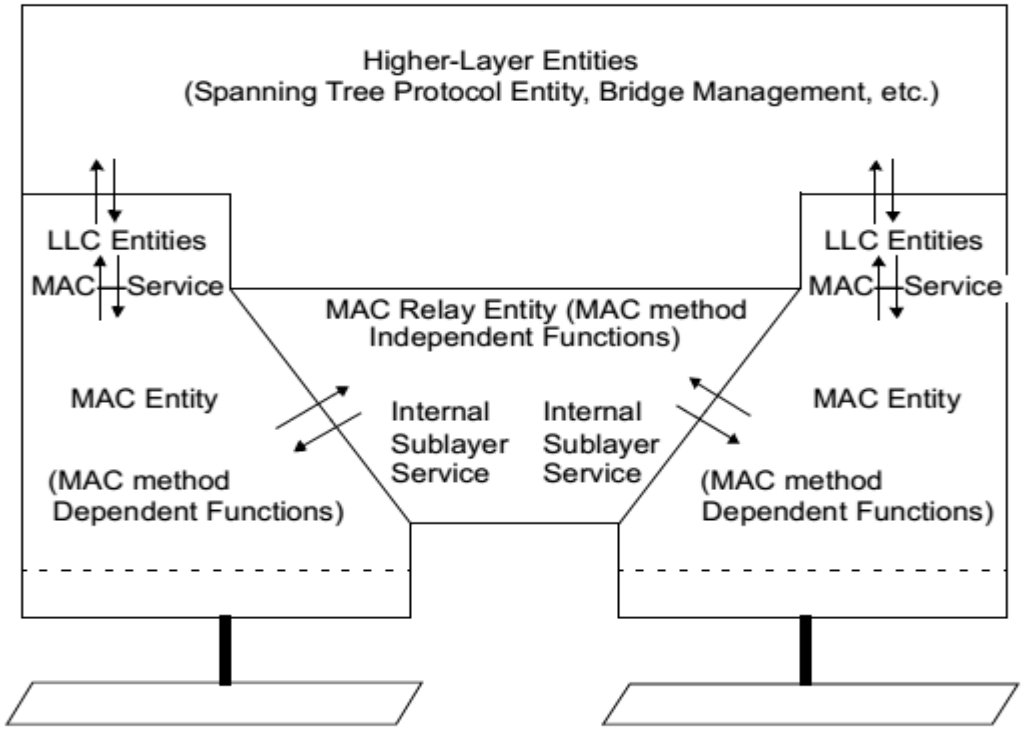
IEEE 802® Reference Model for End Station



IEEE 802.1D Bridge Architecture

M_UNITDATA.indication (
frame_type,
destination_address,
source_address,
mac_service_data_unit,
user_ priority,
frame_check_sequence
)

M_UNITDATA.request (
frame_type,
destination_address,
source_address,
mac_service_data_unit,
user_priority,
access_priority,
frame_check_sequence
)



❑ IEEE 802.1D Bridge shall Conform to IEEE Std 802.2 for the implementation of a class of LLC supporting Type 1 operation as required by 7.3 and 7.1.2. Each Bridge Port shall support the operation of LLC Type 1 procedures in order to support the operation of the Spanning Tree Protocol Entity. Bridge Ports may support other types of LLC procedures, which may be used by other protocols.
❑ A bridge processes protocols in the MAC sublayer and is functionally transparent to LLC and higher layer protocols. MAC frames are forwarded between access domains, or filtered (i.e., not forwarded to certain access domains), on the basis primarily of MAC addressing information

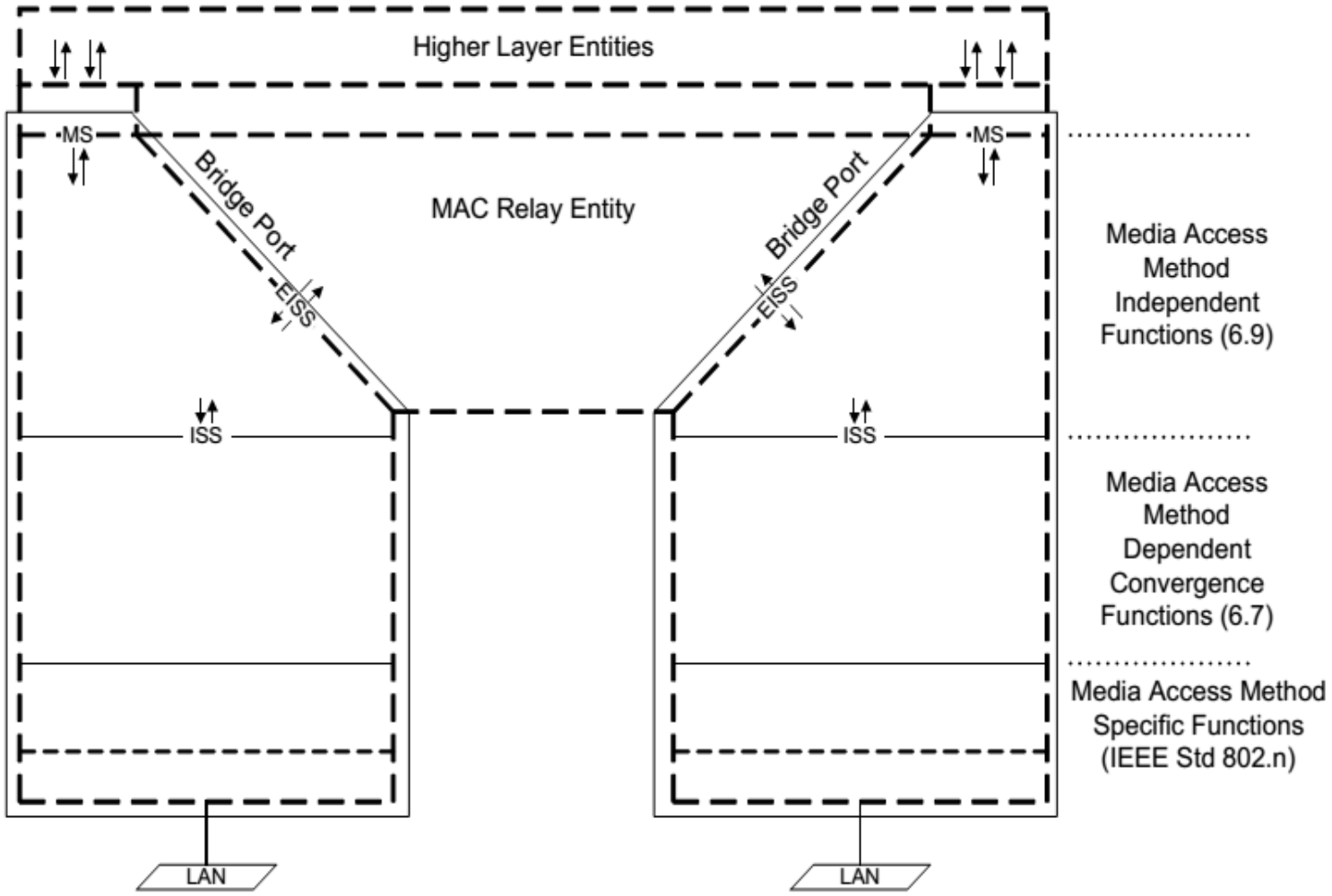
IEEE 802.1Q Bridge Architecture

EM_UNITDATA.indication

(
destination_address,
source_address,
mac_service_data_unit,
priority,
drop_eligible,
vlan_identifier,
frame_check_sequence,
service_access_point_identifier,
connection_identifier,
)

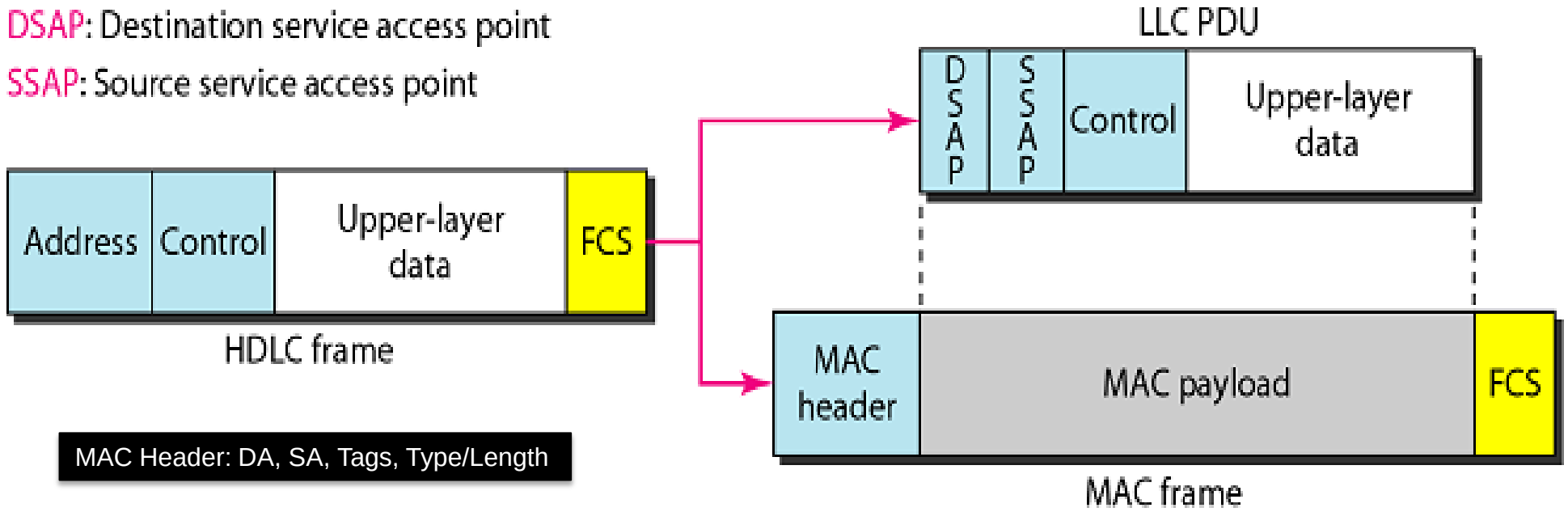
EM_UNITDATA.request

(
destination_address,
source_address,
mac_service_data_unit,
priority,
drop_eligible,
vlan_identifier,
frame_check_sequence,
service_access_point_identifier,
connection_identifier,
)

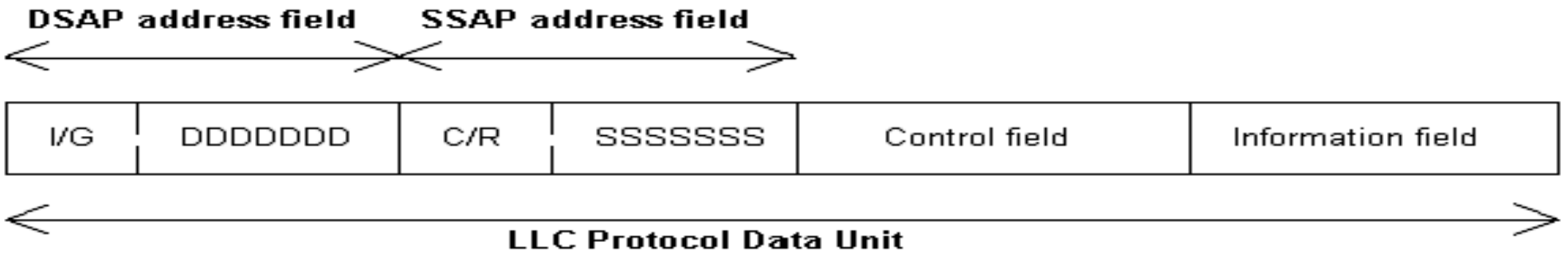


LLC Framing

- The LLC header contains a control field like HDLC and is used for flow and error control.
- The two access point fields (DSAP and SSAP) define the upper layer protocol at the source and destination that uses LLC.



LLC Address Fields



I/G = *Address type designation bit* (indicating individual or group *actual addresses*)

0 = Individual DSAP
1 = Group DSAP

C/R = The *command/response identifier bit*

0 = Command } The value of the C/R bit is set by the
1 = Response } operation of the LLC protocol

D D D D D D D = Destination *actual address*
S S S S S S S = Source *actual address*

Note

i. A complete LLC PDU is shown so that the address fields can be seen in context.
ii. The leftmost bit of each field is the least significant bit.
iii. The Information field is not present in all LLC PDUs

Huawei SDAN Architecture Based on D-CCAP