

Decoupling ECP From EVB

Paul Bottorff

Paul.Bottorff@hpe.com

June 17, 2020

What's The Problem

- ECP is specified in Clause 43 of 802.1Q as reliable link layer transport
- ECP operation is controlled by two state variables specified in clause 43 which are:
 - ackTimerInit= 10s of microseconds to wait for an acknowledgement
 - maxRetries= number of transmission retries before we give up
- Currently the only way provided in 802.1Q to initialize the ECP state variables ackTimerInit and maxRetries is by the Edge Virtual Bridging LLDP TLVs (EVB TLV)
- The EVB LLDP TLV is an application specific TLV used to configure EVB services
 - Both ends of the link send the EVB TLV indicating their roles in the EVB protocols
 - One of the functions of the EVB TLV is to negotiate the ackTimerInit and maxRetries values for ECP
 - The ackTimerInit for ECP is set to the longest time proposed by the local and remote EVB TLVs
 - The maxRetries is set to the largest number proposed by the local and remote EVB TLVs
- The ackTimerInit and maxRetries parameters are reflected in the application specific Edge Virtual Bridging SNMP MIB (EVB MIB) as read only variables
 - In clause 12 the ECP management variables are independent from EVB, though read only
 - In addition, other read only status variables such as transmit and receive counts are specified in for ECP management
- Even though the ECP protocol and clause 12 management variables are EVB independent ECP management is rolled into the SNMP MIB (17.7.20) for EVB and the LLDP EVB MIB (D.5.2).



Strategy To Separate ECP

- Keeping backward compatibility with existing EVB implementations
 - Retain existing EVB SNMP MIB unaltered
 - Maintain existing EVB LLDP TLV unchanged
- Add new ECP management objects and MIB component which includes
 - ecpProposedRTE: Used to set the ackTimerInit state variable
 - ecpProposedR: Used to set the maxRetries state variable
 - Other status variables from the current ECP management variables (12.26.2 table 12-31)



Operation of New ECP management Variables

– Operation of ecpProposedRTE

- If no EVB LLDP TLV is present then the MIB variable ackTimerInit will take the value resulting from ecpProposedRTE. The MIB variable ecpOperAckTimerInit represents ECP ackTimerInit.
- If EVB LLDP TLVs (local and(or) remote) are present then the ackTimerInit will be determined by taking the resulting value from the maximum of the EVB TLV's RTE values (D.2.12.6) and the ecpProposedRTE. The MIB variable ecpOperAckTimerInit will represent ECP ackTimerInit.

– Operation of ecpProposedR

- If no EVB LLDP TLV is present then the value of maxRetries will take the value of the ecpProposedR. The MIB variable ecpOperMaxRetries will represent the value of the ECP variable maxRetries.
- If EVB LLDP TLVs (local and(or) remote) are present then maxRetries will be determined by taking the maximum of the EVB TLV's R values (D.2.12.5) and the ecpProposedR. The MIB variable ecpOperMaxRetries will represent the value of the ECP variable maxRetries.

Changes to clause 12.27: ECP Management

- Clause 12 ECP management is already separate from Edge Virtual Bridge Management (EVB)
- Current ECP management objects are all read only
- We need to add two new R/W objects to table 12-31
 - ecpProposedR: This object provides a value used in a 3 way negotiation with the LLDP EVB TLV R (D.2.12.5) objects (local and remote)
 - ecpProposedRTE: This object provides a value used in a 3 way negotiation with the LLDP EVB TLV RTE (D.2.12.6) objects (local and remote)
- In addition a new ECP MIB component needs to be added which includes the two new management objects as well as the existing status variables.

Changes to clause 43: ECP operation

- Subclause 43.3.6.1 ackTimer last sentence
 - Currently: “This timer is initialized using the value of ackTimerInit determined as stated in D.2.12.6.”
 - Change to: “This timer is initialized with the value of ackTimerInit.”
- New subclause: 43.3.6.2 ackTimerInit
 - Currently: This ECP state machine parameter is not listed even though it is used in the ECP state machine and referenced in clause 12, 43, and annex D.
 - Add subclause: 43.3.6.2 ackTimerInit: This parameter is used to initialize the ackTimer variable. It is set by a three way negotiation between the values of the ecpProposedRTE object of the ECP management database and the local and remote EVB LLDP TLV’s RTE value (D.2.12.6).
- Subclause 43.3.7.4 maxRetries last two sentences
 - “; this variable can be changed by management as documented in 12.26.2. The value is derived from ecpOperMaxRetries (Table 12-31).”
 - These statements are incorrect and inconsistent with annex D (bug in current 802.1Q)
 - Clause 12.26.2 only states that an instance of ECP MIB is generated when the VDP MIB is instantiated.
 - The item in table 12-31 is read only and so can’t be used to set maxRetries. It is only provided to allow a user to determine the outcome of the clause D negotiation.
 - Change to: “The ECP variable maxRetries is set by a three way negotiation between the values of the ecpProposedR object of the ECP management database and the local and remote EVB LLDP TLV’s R value (D.2.12.6).”



Changes to subclause D.2.12: EVB LLDP TLV

– Subclause D.2.12.5

- Currently: “**D.2.12.5 R** This field carries the maxRetries value for the ECP state machine (43.3.7.4). Both sides transmit the local value, and use the largest of the two values of R. If no remote value is available, then the local value is used.”
- Proposed: “**D.2.12.5 R** This field carries a proposed maxRetries value for the ECP state machine (43.3.7.4). The value of maxRetries is the largest of the local EVB R, remote EVB R, and the ECP management object ecpProposedR. The value determined for maxRetries will be reflected in the ECP management object ecpOperMaxRetries.”

– Subclause D.2.12.6

- Currently: “**D.2.12.6 RTE (retransmission exponent)** RTE is an EVB link or S-channel attribute used to calculate the minimum ECPDU retransmission time, ackTimerInit. The value of ackTimerInit, in ECP timer tics of 10 usec, is calculated as $2^{**}RTE$ ECP timer tics. Both sides transmit the local value, and use the largest of the two values of RTE for this calculation. If no remote value is available, then the greater of 2 ms and local value is used.”
- Proposed: “**D.2.12.6 RTE (retransmission exponent)** RTE is an EVB link attribute used to calculate the minimum ECPDU retransmission time, ackTimerInit. The value of ackTimerInit, in ECP timer tics of 10 usec, is calculated as $2^{**}RTE$ ECP timer tics. Both sides transmit their local value, and use the largest of the three values which are the local and remote RTE and the management object ecpProposedRTE. If no remote value is available, then the greater of 2 ms, local value and ecpProposedRTE is used.”



The Aruba logo consists of the word "aruba" in a lowercase, bold, orange sans-serif font. The letters are closely spaced, and the 'a' and 'u' have a distinctive shape with a small gap at the top.

a Hewlett Packard
Enterprise company

Thank you

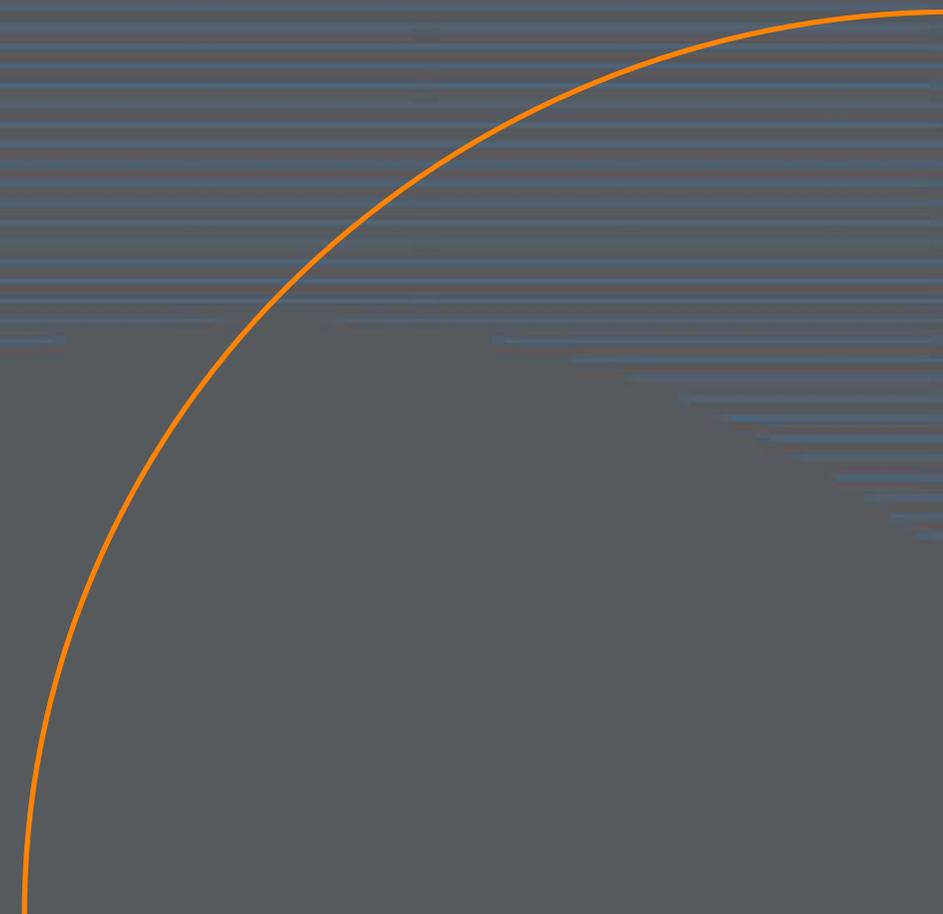
Paul Bottorff
Paul.Bottorff@hpe.com

Backup Slides

aruba

a Hewlett Packard
Enterprise company

ECP Protocol Review



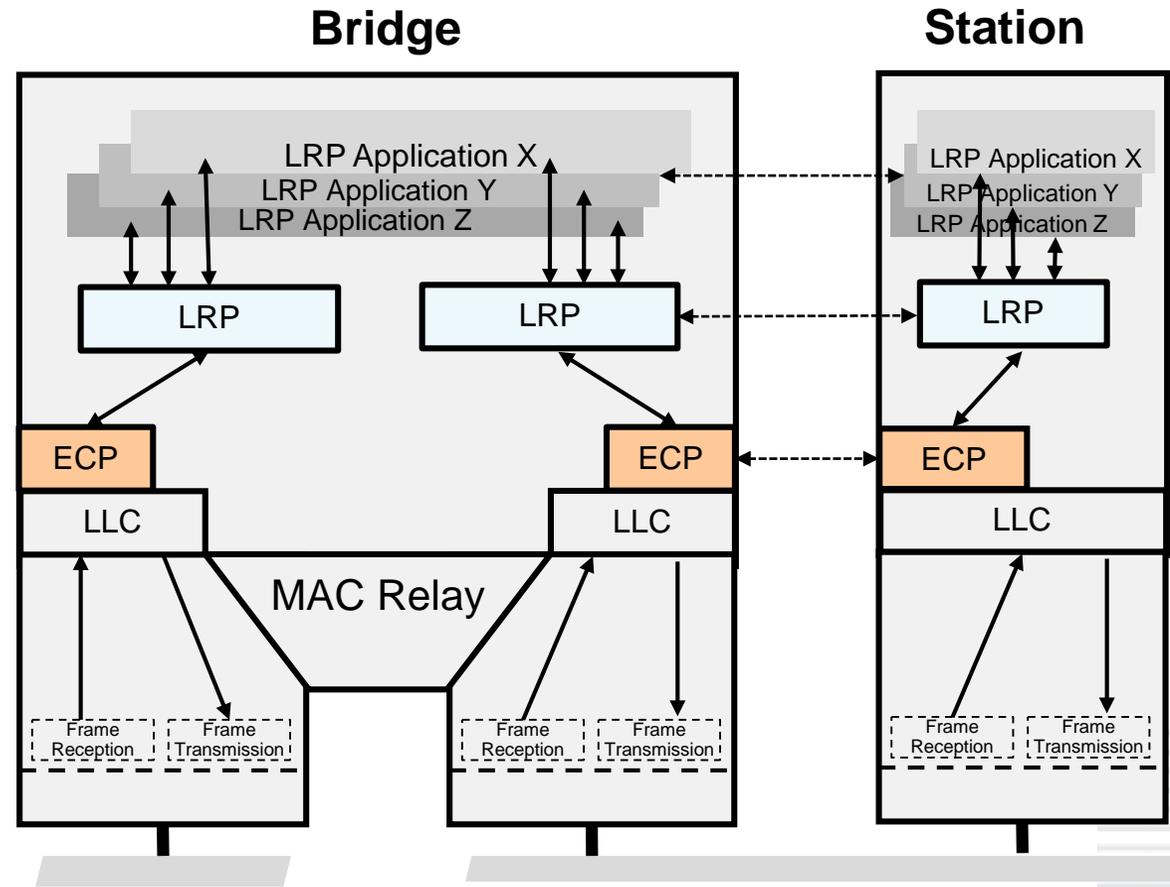
Edge Control Protocol (ECP) Service Characteristics

- ECP as defined in 802.1Q-2014 clause 43 supports transfer of control information between bridges and stations over 802 links
- ECP is currently used by two 802.1 protocols:
 - Virtual Network Interface Discovery and Configuration Protocol (VDP, 802.1Q Clause 41)
 - Port Extender Control and Status Protocol (PE CSP, 802.1BR)
 - Link Local Registration Protocol (LRP, 802.1CS)
- ECP can provide service to multiple Upper Layer Protocols (ULPs)
- Basic services provided by ECP are:
 - Reliable delivery of ULPDUs, resilient against frame loss.
 - In order delivery of ULPDUs to the recipient ULP
 - Delivery of a single copy of each ULPDU to the recipient
 - Flow control that provides protection against buffer overrun and congestion

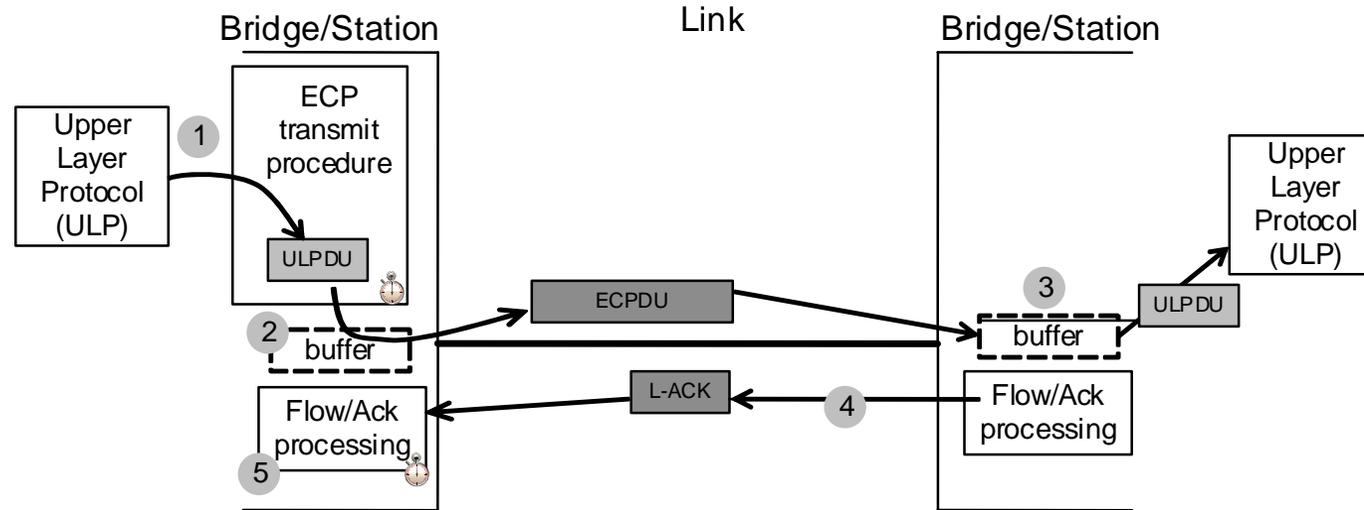


Edge Control Protocol (ECP) in Bridges and End Stations

- ECP is an L2 transport service sitting on top of LLC
- Each ECP dialogs with a peer determined by the destination MAC address used in the ECP frame
- ECP layer provides a service interface for client layers which can be used by any Upper Layer Protocol (i.e. LRP in this example)
- By using ECP rather than LLC for LDP-DT, LRP can be based on a reliable delivery service

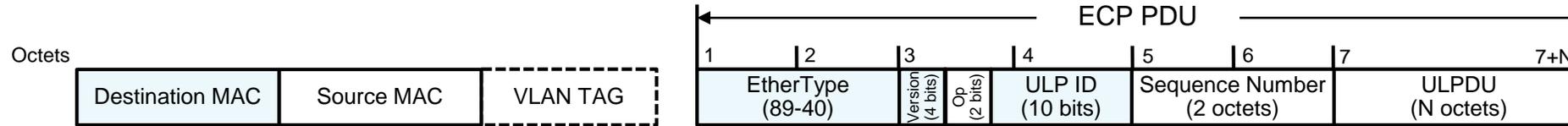


Edge Control Protocol (ECP) Operation



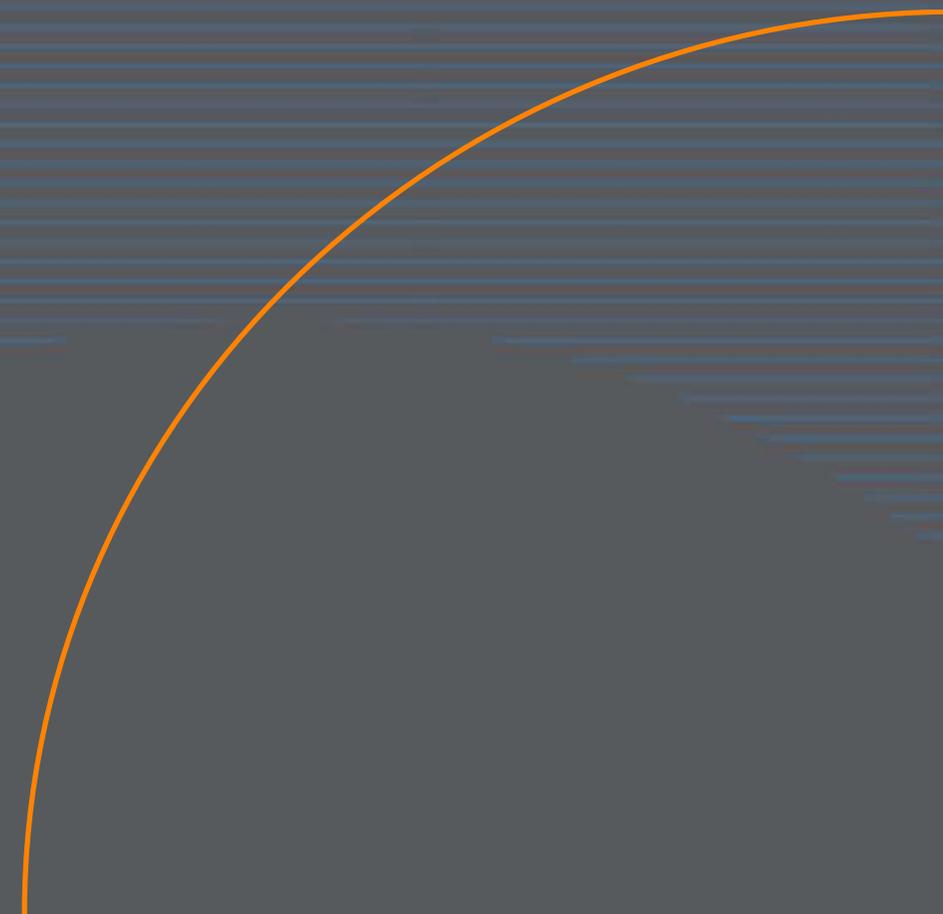
- ECP is a simple stop and wait automatic repeat request protocol supporting frame by frame transport of data from multiple Upper Layer Protocols (i.e. LRP database management protocol).
 - 1 Upper layer (ULP) passes an outgoing ULP Data unit to ECP by invoking a transmit request procedure
 - 2 The ULPDU (perhaps containing a set of TLVs) is transmitted and an ECP low-level acknowledgement (ACK) timer is set. The ULPDU is retained in a local ECP buffer.
 - 3 The ECP frame is received into a receive buffer, here it is held until it is removed by an ECP procedure that passes the ULP Data Unit to the associated ULP.
 - 4 When the receive buffer is emptied, a L-ACK is sent to the sender
 - 5 If the L-ACK is received before the L-ACK timer expires, then the transmit buffer is cleared and ECP can process another ULPDU
 - 5 If the L-ACK timer expires before the L-ACK is received, then the frame in the transmit buffer is re-sent and the L-ACK timer is re-initialized
- Simple enough to allow low level hardware/firmware implementations with reserved frame buffer

Edge Control Protocol (ECP) Frame Format



- ECP frame destinations are addressed by destination MAC + EtherType + Version + Upper Layer Protocol ID
 - Destination MAC addresses may be Group Addresses or Unicast Addresses
 - Currently the VDP protocol uses the Nearest Customer Bridge group address
 - Currently the 802.1BR PE-CSP protocol uses either the Nearest Non-TPMR Bridge group address or a unicast address determined by and LLDP exchange
 - Current VDP for NVO3 uses a unicast ECP address
 - Current LRP protocol uses the Nearest Customer Bridge group address
 - LLC decodes ECP by EtherType 89-40 and Version (0x1)
 - Upper Layer Protocol (ULP) identifiers are encoded in the subtype field
 - Currently ulpID 0x001 identifies the VDP protocol and ulpID 0x002 identifies the 802.1BR PE-CSP protocol
 - IEEE has reserved all values above 3 for future assignment to new protocol (like LRP)
- ECP Operation types are Request and Acknowledgement
- The sequence number is used for acknowledgement and error checking
- The Upper Layer PDU is opaque data passed to the ULP

ECP Configuration Issues



ECP Configuration

- ECP has two configuration parameters which must be set:
 1. maxRetries
 2. ackTimer
- For the Edge Virtual Bridge application (IEEE Std 802.1Q clause 40-42) these parameters are negotiated between the link ends using the LLDP EVB Discovery and Configuration TLV (IEEE Std 802.1Q clause D.2.12)
- ECP can be used without the LLDP by setting the parameters either by some other discovery protocol or by provisioning
- Adding a Configuration and Discovery LLDP TLV for the LRP protocol with ECP configuration parameters is desirable since it allows the negotiation of the ECP parameters between the link ends