Functionality, support, enablement, and management of EMS features

Addressing D1.0 Comments 2, 7, 8, 9, 35, 36, 55, 56, 58, 59, 60, 61, 65, 67, 68, 70, 73, 74

Adee Ran, Cisco

Introduction

- D1.0 defines EMS across multiple sublayers: gRS (Clause 195), PCS (Clause 82/119/133/172/175), RS-FEC (Clause 91/134/161), and Management (Clauses 30 and 45).
- Several comments identify that the draft conflates four distinct concepts:
  - Support / Ability: whether a device or sublayer implements and advertises the optional EMS capability.
  - Enablement: management control that activates the capability (read-write control).
  - Negotiation: some method (e.g., LLDP; currently not specified) confirming peer capability prior to enabling; after negotiation, both peers should have the same enablement.
  - Functionality: any specific behavior (stateless encoding, control ordered sets, RS-FEC error marking) that applies when enabled.
- D1.0 lists different behaviors in each sublayers with a mixture of ability/enable variables and general statements about EMS supports
  - Requirements for 50GBASE-R and 100GBASE-R RS-FEC are found in the PCS clause, while the gRS requirements are separate in clause 195
  - This is architecturally unclean, and difficult to grasp for uninitiated readers
  - It also allows for enablement of a subset of the required features, which can have unintended consequences
- Goal of this presentation: establish a clean architectural definition and management interface, using the 802.3 sublayer stack (RS, PCS, RS-FEC).

Comments addressed

This presentation addresses the following comments on Draft 1.0:

- **2**: Clause 30 enablement
  - The comment suggests that enablement of EMS features should be conditional on verifying link partner support. This means enablement is an all-or-none operation, a fundamental concept for many other comments.
  - This comment is in the bucket with proposed AIP. This presentation does not suggest changes to the proposed response.
- **7, 35, 36, 55**: PCS requirements & scope
- **60, 68, 70**: RS-FEC requirements
- **8, 9, 56, 58, 59, 61, 65, 67, 73, 74** : PICS Alignment

Other comments may also be addressed if the proposed changes are accepted.

EMS support - principles

- EMS requires specific behaviors in multiple sublayers
  - Sublayers should be considered as separate pieces of IP (even if some implementations integrate some of them together)
  - In a Physical Layer implementation, all sublayers need to **support** the required behaviors
  - We typically define ability management variables in each affected clause
- If an EMS capability is supported by all sublayers, it can be **enabled**
  - Enabling only part of the behaviors does not achieve the required functionality (and thus is not in the scope of the project); it might break some assumptions
  - If an EMS capability is enabled, all sublayers should enable all required behaviors
  - If an EMS capability is not enabled, all sublayers should keep the existing standard behaviors
  - Enablement should be the same in all sublayers; a single “enable” management variable achieves that
- Common configuration among link partners
  - EMS is symmetric; both partners must have the same enablement setting
  - Negotiation of support and activation are currently defined as out of scope

Clause 82 (PCS) Functionality & Sublayer Modularity

- Issues in 82.5a
  - Layering violation: specifies RS-FEC codeword error marking requirements in the PCS clause. RS-FEC is an independent sublayer (Clause 91/161) and CAUI-n is separate annexes; their requirements should reside in the corresponding clauses and annexes.
  - Unconditional “shall” statements: requirements are stated as if “support” mandates using the stateless encoding and control ordered set insertion, regardless of enablement.
- Multiple/missing management variables
  - PCS features in 82.2.4 (transmit) and 82.2.17 (receive) are controlled by individual control variables (PCS_64B/66B_stateless_enable, BER_monitor_disable)
  - These should be enabled together when the EMS link feature is enabled in the PCS (along with other sublayers)
  - If EMS link is not enabled, the state-diagram encoder/decoder must be used (because FEC correction is not guaranteed), control ordered sets must not be sent, and should be treated as error if received
- State-diagram encoder and decoder are redefined to be optional
  - But with non-EMS PHYs, e.g. with no RS-FEC or with correction bypassed, the state diagram serves an important role (recall MTTFPA)
  - This amendment should not modify existing specifications beyond its scope
- Other PCS clauses have similar issues

Clause 82 (PCS) Functionality & Sublayer Modularity

- Proposed solution:
  1. Define two support/enablement variable pairs, ems_link_ability+ ems_link_enable, ems_packet_ability+ems_packet_enable (which would be common to all subclauses that define EMS functionality)
  2. Restructure 82.5a strictly for PCS functionality:
     1. When ems_packet_enable is TRUE: BER monitor shall be disabled
     2. When ems_link_enable is TRUE: BER monitor shall be disabled, shall use stateless encoder (82.2.4.2) and stateless decoder (82.2.17.2), shall encode and decode CtlOS (82.2.3.9)
     3. When both are FALSE: BER monitor shall be enabled, shall use state-diagram encoder and decoder
  3. Relocate RS-FEC requirements out of Clause 82 into Clauses 91, 134, and 161.
- Similar in other PCS clauses

Proposed text in 82.5a

82.5a Ethernet Metadata Services (EMS) Functionality

A 100GBASE-R PCS may support Ethernet Metadata Services (EMS) specified in Clause 195. Support of link metadata is indicated by the variable ems_link_ability. Support of packet metadata is indicated by the variable ems_packet_ability.

If a PCS supports link metadata or packet metadata, enablement of these features is controlled by the variables ems_link_enable and ems_packet_enable, respectively. These variables have no effect in a PCS that does not support the corresponding features.

NOTE 1—Enablement of EMS capabilities in a Physical Layer affects other sublayers as well as the PCS.

When ems_link_enable is TRUE, the PCS shall use the stateless encoder (82.2.4.2) and the stateless decoder (82.2.17.2), and shall enable the transmission and reception of control ordered sets as defined in 82.2.3.9.

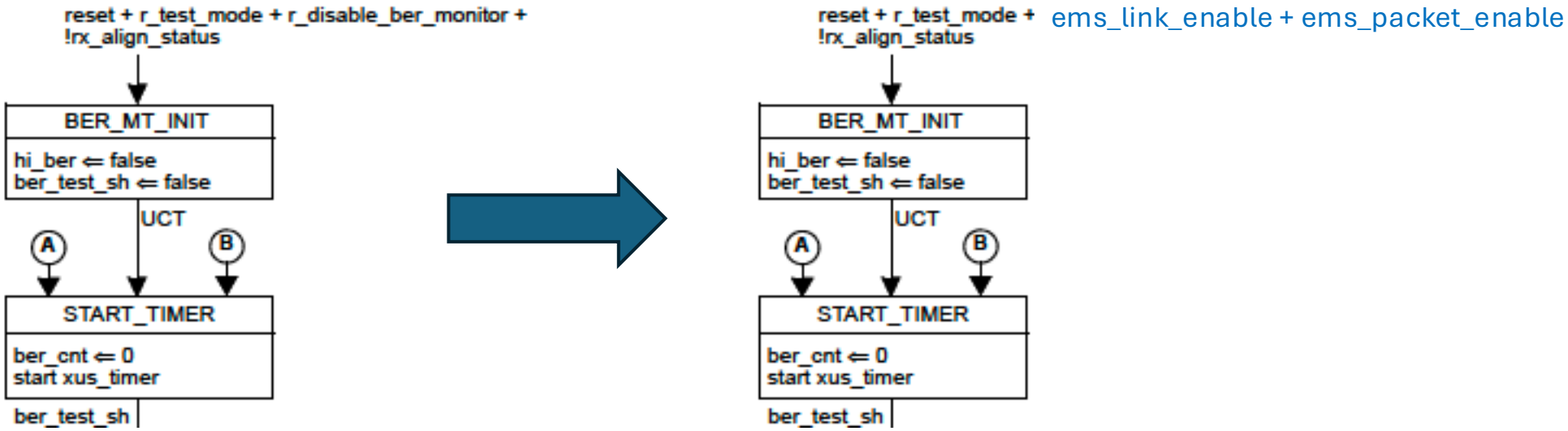
When ems_link_enable is FALSE, the PCS shall use the state diagram encoder (82.2.4.1) and the state diagram decoder (82.2.17.1), and shall classify control ordered sets as errors.

NOTE 2—When either ems_link_enable or ems_packet_enable is TRUE, the BER monitor state diagram (Figure 82–15) is always in the BER_MT_INIT state, and effectively disabled.

Delete D1.0 text related to the RS-FEC and CAUI-n – out of scope for the PCS

Proposed change in Figure 82–15

(only top arrow condition changes)



Other changes in Clause 82

- In 82.2.4, change the second paragraph (in D1.0) to

  For 100GBASE-R, the PCS shall include the state-diagram encoder (82.2.4.1) and may include the stateless encoder (82.2.4.2). If ems_link_ability and ems_link_enable are both TRUE, the stateless encoder shall be used. Otherwise, the state-diagram encoder shall be used.

- In 82.2.17, change the third paragraph (in D1.0) to

  For 100GBASE-R, the PCS shall include the state-diagram decoder (82.2.17.1) and may include the stateless decoder (82.2.17.2). If ems_link_ability and ems_link_enable are both TRUE, the stateless decoder shall be used. Otherwise, the state-diagram decoder shall be used.

- In 82.3.1, add new variables ems_link_ability, ems_packet_ability, ems_link_enable, ems_packet_enable
- Delete the new variables r_disable_ber_monitor (82.2.19.2), BER_monitor_disable, PCS_64B/66B_stateless_enable, PCS_64B/66B_stateless_ability, BER_monitor_disable_ability, and control_ordered_set_ability (82.3.1)
  - Delete references to these variables (replace with the new “enable” variables as appropriate)
- Update the PICS as necessary (e.g., as suggested by comments 8 & 9).

Proposed changes in Clause 91

- In 91.6, replace the “local” FEC_full_error_marking variables with the “global” variables ems_link_ability, ems_packet_ability, ems_link_enable, ems_packet_enable
- Add a new subclause similar to 82.5a:

  91.5a Ethernet Metadata Services (EMS) Functionality

  An RS-FEC sublayer may support Ethernet Metadata Services (EMS) specified in Clause 195. Support of link metadata is indicated by the variable ems_link_ability. Support of packet metadata is indicated by the variable ems_packet_ability.

  If an RS-FEC supports link metadata or packet metadata, enablement of these features is controlled by the variables ems_link_enable and ems_packet_enable, respectively. These variables have no effect in an RS-FEC that does not support the corresponding features.

  NOTE—Enablement of EMS capabilities in a Physical Layer affects other sublayers as well as the RS-FEC.

  When ems_link_enable and ems_packet_enable are both FALSE, the RS-FEC Reed-Solomon Decoder (see 91.5.3.3) shall use the sparse error marking method.

  When either ems_link_enable or ems_packet_enable is TRUE, the RS-FEC Reed-Solomon Decoder shall use the full error marking method, and shall correct and indicate errors, treating FEC_bypass_correction_enable and FEC_bypass_indication_enable as false.

- Update the PICS as necessary.

Proposed changes in Clause 195

- Insert a new subclause before 195.4 (currently labeled 195.3a) that lists the functionality required from the different sublayers, and defines the ability and enable variables, which are common to all sublayers.
  - This clause should also specify that the enablement of an EMS feature shall be done only after negotiation with the link partner, such that both partners have the same enablement state.
- See proposed text in the next slide.

Proposed text for new 195.3a

195.3a Ethernet Metadata Services sublayer support and enablement

EMS behavior spans multiple sublayers. Support of EMS link metadata or packet metadata in a Physical Layer requires support in all sublayers for which EMS behavior is defined. Similarly, enablement of EMS link metadata or packet metadata in a Physical Layer is common to all sublayers for which EMS behavior is defined.

If EMS is supported by a Physical Layer, it shall be disabled by default and whenever communication with the link partner is not established. When communication with the link partner has been established, EMS will only be enabled after verification that the link partner supports the same metadata capability and the associated syntax and semantics of the metadata to be exchanged. The mechanism by which such verification is conducted is beyond the scope of this standard.

Proposed text for new 195.3a (cont.)

195.3a.1 Management variables

The following variables are defined in each sublayer for which EMS behavior is specified. Each of these variables shall have the same value in all sublayers.

ems_link_ability

Boolean variable indicating whether the sublayer supports the link metadata capability. The value TRUE indicates that link metadata is supported.

ems_packet_ability

Boolean variable indicating whether the sublayer supports the packet metadata capability. The value TRUE indicates that packet metadata is supported.

ems_link_enable

Boolean variable that controls the activation of link metadata functionality in the sublayer. When set to TRUE, link metadata operation is enabled. When set to FALSE, standard non-EMS operation is used. The default value is FALSE.

ems_packet_enable

Boolean variable that controls the activation of packet metadata functionality in the sublayer. When set to TRUE, packet metadata operation is enabled. When set to FALSE, standard non-EMS operation is used. The default value is FALSE.

Add standard MDIO mapping text and table

Proposed text for new 195.3a (cont.)

195.3a.2 EMS requirements

The requirements for each sublayer when EMS is supported and enabled are summarized as follows

- gRS: When packet metadata is enabled (ems_packet_enable = TRUE), the gRS supports the transfer of packet metadata between the EMS client and PCS as specified in this Clause. When link metadata is enabled (ems_link_enable = TRUE), the gRS supports the transfer of link metadata between the EMS client and the PCS as specified in this Clause.
- PCS: The requirements for PCS behavior when link metadata or packet metadata are enabled are specified in 133.5a for 50GBASE-R, in 82.5a for 100GBASE-R, in 119.5a for 200GBASE-R and 400GBASE-R, and in 172.5a for 800GBASE-R and 1.6TBASE-R.
- FEC: 50GBASE Physical Layers are required to include the RS-FEC sublayer (clause 134) in the same package as the PCS (no LAUI-2 in between). 100GBASE Physical Layers are required to include either the RS-FEC (clause 91) or RS-FEC-Int (Clause 161) sublayer in the same package as the PCS (no CAUI-4 in between). The requirements for FEC sublayer behavior when link metadata or packet metadata are enabled are specified in 134.5a for 50GBASE-R RS-FEC, in 91.5a for 100GBASE-R RS-FEC, and in 161.5a for 100GBASE-R RS-FEC-Int.

Changes in other clauses

- Clause 119
  - Update 119.5a similar to the proposed changes in 82.5a, but include the statement “FEC bypass indication mode shall not be enabled”
  - Add the same ability/enable variables
  - Update PICS
- Clause 133: updates similar to those proposed for clause 82
- Clause 134: updates similar to those proposed for clause 91
- Clause 161: updates similar to those proposed for clause 91
- Clause 172: updates similar to those proposed for clause 119
- Clause 175: updates similar to those proposed for clause 119

Summary

- This presentation proposes a set of changes to achieve:
  - Clear taxonomy: distinguish ability, enablement, negotiation (peer confirmation), and operation (normative behaviors when enabled).
  - Clean sublayer separation: each clause has specific EMS requirements.
  - Clean management specifications: use common variables **ems_link_ability**, **ems_link_enable**, **ems_packet_ability**, **ems_packet_enable** in all clauses, and clarify the requirement to enable only after link partner support has been established.
- Recommendation: implement the proposed changes in slides 7-14 with editorial license, to resolve comments 2, 7, 8, 9, 56, 58, 60, 61, 65, 67, 68, 70, 73, and 74 (and possibly other comments).

That’s all!

Questions?