IEEE P802.3by
Auto-Negotiation of FEC

reference P802.3 Draft 1.0 Comment #62

Dan Dove, DNS for Avago Technologies

May Interim Meeting, Pittsburgh 2015

May 12, 2015

Supporters

- TBD

Introduction

- The current draft if IEEE p802.3by (d1p0) states

73.6.5 FEC capability

FEC (F2:F3:F0:F1) is encoded in bits D446:D47 of the base link codeword. The four ~~two~~ FEC bits are used as follows:

a) F0 is FEC ability
b) F1 is FEC requested
c) F2 is 25G RS-FEC requested
d) F3 is 25G BASE-R FEC requested

Bits F2 and F3 are used for resolving FEC operation for 25G PHYs while bits F0 and F1 are used for other speeds of operation. Bits F0 and F1 are not used for 25G PHYs.

For 25G PHYs if neither PHY requests FEC operation in bits F2 or F3 then FEC is not enabled.

For 25GBASE-KR and 25GBASE-CR PHYs if either PHY requests RS-FEC then RS-FEC operation is enabled, otherwise if either PHY requests BASE-R FEC then BASE-R operation is enabled.

For 25GBASE-KR-S and 25GBASE-CR-S PHYs if either PHY requests RS-FEC or BASE-R FEC then BASE-R operation is enabled. This is because 25GBASE-KR-S and 25GBASE-CR-S PHYs do not support RS-FEC operation.

Problem Statement & Proposal

Bits F2 and F3 are used for resolving FEC operation for 25G PHYs while bits F0 and F1 are used for other speeds of operation. Bits F0 and F1 are not used for 25G PHYs.

There is no clear rationale for removing use of F0 to indicate FEC ability for 25G operation. This bit should behave the same regardless of speed.

F1 already enables the use of FEC. There is no need to change its use. The question occurs, "Why ignore existing bits, and add two bits to decide which 25G FEC to use?

With discussion, I came to the conclusion that the original spec was "non-obvious" on what F1 actually does.

While it overtly indicates "FEC Requested", with addition of the current text it doesn't clearly articulate which kind of FEC is to be requested.

Is it requesting 10Gb/lane FEC?
Is it requesting 25Gb/lane FEC?

The additional text is intended to eliminate that question and thus creates a new bit (F2) to indicate RS-FEC requested and another new bit (F3) to indicate 25G BASE-R FEC requested. These two bits are required because 25GBASE-CR supports two different modes of FEC, which makes it a novel PHY in that regard.

The two additional bits allow one to Auto-Negotiate in a single-pass between 10G w/o FEC and 25G w/FEC for instance. But its not the obvious way to address that problem, nor is it neatly scalable.

Proposal:

1. **Rename F1 to "10Gb/lane FEC Requested"**
2. **Rename F2 to "25Gb/lane FEC Requested"**
3. **Remove BASE-R FEC from 25GBASE-CR capabilities**

Rationale

- Proposed new text simplifies the change to clause 73
  - While simplification of changes is not necessarily a goal, it reduces risk of extraneous and unintended growth in the standard
- Proposed new text is logically consistent with the existing standard
  - Prior specifications, for 10G, 40G and 100G perform the same functionality using F0 and F1 despite different FEC approaches being used among them
- Proposed new text minimizes the number of bits required to obtain desired functionality
  - F2 communicates the desire to use 25Gb/lane FEC and we decide which type of FEC via Auto-Negotiation of the PHY type.
  - If a 25GBASE-KR/CR-S PHY, F0 is set, F1 may be set (if 10G FEC req) and F2 will determine whether BASE-R FEC is requested.
  - If a 25GBASE-KR/CR PHY, F0 is set, F1 may be set (if 10G FEC req) and F2 will determine whether RS-FEC FEC is requested
- Proposed new text is scalable for future PHYs
  - Unless a future PHY defines more than one FEC type at 25Gb/lane, there will be no need to modify text.
  - If a future 50Gb/lane or 100Gb/lane specification comes out, they can simply take additional bits but retain the current logical extension.
    - For instance;
      - F3 would be “50Gb/lane FEC Requested”
      - F4 would be “100Gb/lane FEC Requested”

Rationale (continued)

- Behavior of devices should be defined by their PHY type.
  - CR/KR-S PHY specification should only define use of BASE-R FEC
  - CR/KR PHY specification should only define use of RS-FEC.
  - With this approach, the application space is supported exactly as-is. ***There is no real benefit to having a PHY type with two modes of FEC***.
    - It adds complexity for no benefit.
    - It creates a unique two-FEC PHY mandating unique treatment during AN
- Implementations would behave in a similar manner
  - If host determines BASE-R FEC is desired, it could remove 25GBASE-CR PHY advertisement and set F0 and F2.
  - If host determines RS-FEC is desired, it could advertise 25GBASE-CR and and set F0 and F2.
- A PHY designed to the proposed new text would appear functionally identical to a PHY designed to the existing text.
  - A 25GBASE-CR-S PHY would not have RS-FEC
  - While its *possible* somebody would build a 25GBASE-CR PHY w/o BASE-R FEC within the die, its unlikely for the same reason argued in prior meetings. The die cost relative to overall IC is negligible.
  - While the specification would show two different PHY types, the implementation of 25GBASE-CR will be (with unlikely exception) a single architecture that supports either RS-FEC or BASE-R FEC.
    - Other than AN bits, physically identical.

Logic Table

PHY Type (1)	F0*F0	F1||F1	F2||F2	Outcome
Existing PHY Types (inc 25Gb/lane)	0	NA	NA	FEC ability not present, No FEC
Existing PHY Types (10Gb/lane)	1	0	NA	10Gb/lane FEC not requested by either partner, No FEC
Existing PHY Types (10Gb/lane)	1	1	NA	10Gb/lane FEC enabled
25GBASE-CR-S 25GBASE-KR-S	1	NA	0	25Gb/lane FEC not requested by either partner, no FEC
25GBASE-CR-S 25GBASE-KR-S	1	NA	1	25Gb/lane FEC requested by either partner, BASE-R FEC enabled
25GBASE-CR	1	NA	0	25Gb/lane FEC not requested by either partner, no FEC
25GBASE-CR	1	NA	1	25Gb/lane FEC requested by either partner, RS-FEC enabled

(1) – The PHY Type that has been resolved via Auto-Negotiation. (ie: HCD)

Explicit Changes

Changes to Clause 73:

73.6.5 FEC capability

FEC (F2:F0:F1) is encoded in bits D45:D47 of the base link codeword.
The three FEC bits are used as follows:

a) F0 is FEC ability
b) F1 is 10Gb/lane FEC requested
c) F2 is 25Gb/lane FEC requested

Bit F0 communicates whether the PHY has FEC capability to offer. Bit F1 is used for resolving FEC operation for 10Gb/lane PHYs while bit F2 is used for resolving FEC operation for 25Gb/lane PHYs.

If neither PHY requests FEC operation in bits F1 or F2 then FEC is not enabled.

When the FEC ability bit F0 is set to logical one, it indicates that the PHY has FEC ability (see Clause 74). When the 10Gb/lane FEC requested bit F1 or the 25Gb/lane FEC requested bit F2 are set to logical one, that indicates a request to enable FEC on the link.

Since the local device and the link partner may have set the FEC capability bits differently, the priority resolution function is used to enable FEC in the respective PHYs.

The FEC function shall be enabled on the link if 10GBASE-KR, 40GBASE-KR4, 40GBASE-CR4, or 100GBASE-CR10 is the HCD technology (see 73.7.6), both devices advertise FEC ability on the F0 bits, and at least one device requests FEC on the F1 bits; otherwise FEC shall not be enabled.

The FEC function shall be enabled on the link if 25GBASE-KR, 25GBASE-CR, 25GBASE-KR-S or 25GBASE-CR-S is the HCD technology (see 73.7.6), both devices advertise FEC ability on the F0 bits, and at least one device requests FEC on the F2 bits; otherwise FEC shall not be enabled.

Explicit Changes

Changes to Clause 69:

Table 69–1a—Nomenclature and clause correlation for 25 Gb/s Backplane Ethernet Physical Layers

Nomenclature	Clause										
	73	74	78	106		107	108	109	109A	111	
	AN	BASE-R FEC	EEE	RS	25G-MII	25GBASE-R PCS	RS-FEC	25GBASE-R PMA	25G-AUI C2C	25GGBASE-KR PMD	25GGBASE-KR-S PMD
25GBASE-KR	M	M	O	M[a]	O[a]	M	M	M	O	M	
25GBASE-KR-S	M	M	O	M	O	M		M	O		M

[a]O = Optional, M = Mandatory

leave blank

Explicit Changes

Changes to Clause 74:

Pg 62 Line 29-32: Strike 25GBASE-CR, 25GBASE-KR and Clause 110 references

The 25GBASE-CR, 25GBASE-CR-S, 25GBASE-KR and 25GBASE-KR-S PHYs described in Clause 110 and Clause 111 are required to implement the FEC sublayer and may use it with links with a BER of 10^{-8} or better.

Pg 71 Line 21, 23: Strike clause 74 from 25GBASE-KR and 25GBASE-CR rows.

25GBASE-KR	74, 107, 108, 109, 111
25GBASE-KR-S	74, 107, 109, 111
25GBASE-CR	74, 107, 108, 109, 110

Changes to Clause 105:

Pg 77 Line 42, 47: Change as follows.

25 Gb/s PHY equivalent to 25GBASE-CR ~~without support for the~~ with support for the BASE-R FEC sublayer rather than the RS-FEC sublayer (see Clause 110).

…

25 Gb/s PHY equivalent to 25GBASE-KR ~~without support for the~~ with support for the BASE-R FEC sublayer rather than the RS-FEC sublayer (see Clause 110).

Explicit Changes

Changes to Clause 105:

Table 105–2—Nomenclature and clause correlation, 25GBASE-R

Nomenclature	Clause/Annex													
	73	74	78	106		107	108	109	109A	110		111		112
	Auto-Negotiation	BASE-R FEC	EEE	RS	25G-MII	25GBASE-R PCS	RS-FEC	PMA	25G-AUI C2C	25GBASE-CR PMD	25GBASE-CR-S PMD	25GBASE-KR PMD	25GBASE-KR-S PMD	25GBASE-SR PMD
25GBASE-CR	M	M	O	M	O	M	M	M	O	M				
25GBASE-CR-S	M	M	O	M	O	M		M	O		M			
25GBASE-KR	M	M	O	M	O	M	M	M	O			M		
25GBASE-KR-S	M	M	O	M	O	M		M	O				M	
25GBASE-SR			O	M	O	M	M	M	O					M

leave blank

Explicit Changes

Changes to Clause 110:

Change to N/A

Table 110–1—Physical Layer clauses associated with the 25GBASE-CR and 25GBASE-CR-S PMDs

Associated clause	25GBASE-CR	25GBASE-CR-S
106—RS	Required	Required
106—25G-MII[a]	Optional	Optional
107—PCS	Required	Required
74—BASE-R FEC[b]	Required	Required
108—RS-FEC[b]	Required	N/A
109—PMA	Required	Required
109A—25G-AUI C2C	Optional	Optional
73—Auto-Negotiation	Required	Required
78—Energy Efficient Ethernet	Optional	Optional

Explicit Changes

Changes to Clause 110:

110.6 FEC modes

A 25GBASE-CR PHY implements ~~the BASE-R FEC sublayer (Clause 74) and~~ the 25GBASE-R RS-FEC sublayer (Clause 108). A 25GBASE-CR-S PHY implements the BASE-R FEC sublayer (Clause 74). Each FEC sublayer can be either enabled or disabled, according to AN resolution or management control.

Three FEC modes are supported:

a) When the 25GBASE-R RS-FEC sublayer is enabled, the PHY is defined to operate in the RS-FEC mode.

b) When the BASE-R FEC sublayer is enabled, the PHY is defined to operate in the BASE-R FEC mode.

c) When no FEC sublayer is enabled, the PHY is defined to operate in the no-FEC mode.

A 25GBASE-CR PHY can operate in RS-FEC, ~~BASE-R FEC~~ or no-FEC mode. A 25GBASE-CR-S PHY can operate in either BASE-R FEC or no-FEC mode.

The cable assembly types (CA-N, CA-S or CA-L, see 110.10) that the PHY supports and the required PMD receiver characteristics (110.8.4) depend on the FEC mode.

The FEC mode is determined using AN (Clause 73) and is used in both transmit direction and receive direction. It is recommended to configure the AN FEC advertisement such that only modes that are compatible with the type of the cable assembly attached to the MDI are selected.

Explicit Changes

Changes to Clause 110:

110.8.4.2 Receiver interference tolerance test

Receiver interference tolerance is measured according to the requirements listed in 110.8.4.2.1 through 110.8.4.2.5.

A 25GBASE-CR PHY shall comply with the receiver interference tolerance test requirements for the RS-FEC,~~|BASE-R FEC~~ and no-FEC modes. A 25GBASE-CR-S PHY shall comply with the receiver interference tolerance test requirements for the BASE-R FEC and no-FEC modes.

Table 110–7—~~25GBASE-CR and~~ 25GBASE-CR-S interference tolerance parameters, no-FEC mode

Parameter	Test 1 (low loss)	Test 2 (high loss)	Units

Proposed Motion

Editors adopt recommendation provided in dove_3by_01_0515 to do the following;

1. Rename F1 to “10Gb/lane FEC Requested”
2. Rename F2 to “25Gb/lane FEC Requested”
3. Remove BASE-R FEC from 25GBASE-CR capabilities
4. Make edits shown in slides 7-10 with editorial discretion to modify proposed changes and/or make additional changes to fulfill items 1-3 above.

Q&A