FUJITSU

shaping tomorrow with you

Single 25GBASE-CR PMD with optional RS-FEC

Yasuo Hidaka
Fujitsu Laboratories of America, Inc.

October 7, 2015

Introduction

FUJITSU

- 2 PMDs and 3 FEC modes are very confusing
- 2 PMDs are not necessary, because their difference is merely availability of RS-FEC, and analog frontend is most likely same
- FEC resolution is unnecessarily complicated by use of 2 PMDs
  - This is because arbitration process to choose FEC mode is split between
    1. Priority-based resolution to find HCD (Highest Common Denominator) of PMD
    2. Logic-based resolution to resolve various demands of FEC usages
  - A unified logic-based resolution for single PMD is equivalent and simpler
- Single PMD with optional RS-FEC should be sufficient and significantly reduces confusion

Draft D2.1, P55: FEC capability and resolution

FUJITSU

73.6.5 FEC capability

FEC (F2:F3:F0:F1) is encoded in bits D4~~4~~6:D47 of the base link codeword. The four ~~two~~ FEC bits are used as follows:

a) F0 is 10 Gb/s per lane FEC ability
b) F1 is 10 Gb/s per lane 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 10 Gb/s per lane operation. Bits F0 and F1 are not used for 25G PHYs.

Insert new subclause "73.6.5.1 FEC resolution for 25G PHYs" before second paragraph of 73.6.5 as follows:

73.6.5.1 FEC resolution 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.

Draft D2.1, P54 & P57: Tech Ability and Priority

FUJITSU

Table 73–4—Technology Ability Field encoding

Bit	Technology
A0	1000BASE-KX
A1	10GBASE-KX4
A2	10GBASE-KR
A3	40GBASE-KR4
A4	40GBASE-CR4
A5	100GBASE-CR10
A6	100GBASE-KP4
A7	100GBASE-KR4
A8	100GBASE-CR4
A9	25GBASE-KR-S or 25GBASE-CR-S
A10	25GBASE-KR or 25GBASE-CR
~~A9 through A24~~ A11 through A22	Reserved for future technology

A10:RS-FEC ability →

Table 73–5—Priority Resolution

Higher priority →

Lower priority →

7	25GBASE-KR or 25GBASE-CR	25 Gb/s 1 lane
8	25GBASE-KR-S or 25GBASE-CR-S	25 Gb/s 1 lane, short reach
~~7~~9	10GBASE-KR	10 Gb/s 1 lane
~~8~~10	10GBASE-KX4	10 Gb/s 4 lane
~~9~~11	1000BASE-KX	1 Gb/s 1 lane, lowest priority

dudek_3by_03a_0315.pdf(slide 4) vs D2.1

FUJITSU

Table 73–5—Priority Resolution

7	25GBASE-KR or 25GBASE-CR	25 Gb/s 1 lane
8	25GBASE-KR-S or 25GBASE-CR-S	25 Gb/s 1 lane, short reach

A9	25GBASE-KR-S or 25GBASE-CR-S
A10	25GBASE-KR or 25GBASE-CR

a) F0 is 10 Gb/s per lane FEC ability
b) F1 is 10 Gb/s per lane FEC requested
c) F2 is 25G RS-FEC requested
d) F3 is 25G BASE-R FEC requested

A10 (Remote) AND A10 (Local)

HCD resolved between Two PMDs with CR being higher priority than CR-S

F2 (Remote) OR F2 (Local)

F3 (Remote) OR F3 (Local)

	HCD		OR of RS-FEC REQ	OR of 25G BASE-R REQ	Usage
	CR	CR-S			
0		Y	0	0	noFEC
0		Y	X	1	BASE-R
0		Y	1	X	BASE-R
1	Y		0	0	noFEC
1	Y		0	1	BASE-R
1	Y		1	X	RSFEC

HCD is Highest Common Demoninator with CR being higher priority than CR-S
OR is the OR function of the appropriate FEC Request from the two ends.
X is don't care.

Unified Logic-Based Resolution w/ Single PMD

FUJITSU

- Define single PMD with optional RS-FEC
  - F2 = 25G RS-FEC ability
  - F3 = 25G RS-FEC requested
  - F4 = 25G BASE-R FEC requested

F2 (Remote) AND F2 (Local)

Single PMD (No need to resolve)

F3 (Remote) OR F3 (Local)

F4 (Remote) OR F4 (Local)

	HCD		OR of RS-FEC REQ	OR of 25G BASE-R REQ	Usage
	CR	CR-S			
0		Y	0	0	noFEC
0		Y	X	1	BASE-R
0		Y	1	X	BASE-R
1	Y		0	0	noFEC
1	Y		0	1	BASE-R
1	Y		1	X	RSFEC

HCD is Highest Common Demoninator with CR being higher priority than CR-S
OR is the OR function of the appropriate FEC Request from the two ends.
X is don't care.

A Special Case to Lose Small Flexibility

FUJITSU

- In most cases, we do not lose any flexibility
- However, there is a special case to lose small flexibility
  - By requesting RS-FEC and *deliberately not advertising A9* (CR-S ability) but only A10 (CR ability), we may avoid link up with CR-S PMD in BASE-R FEC mode and allow link up only with CR PMD in RS-FEC mode
    - If we do this and link partner is CR-S, link does not come up (there is no HCD)
    - This is *non-standard* usage of A9, because CR should advertised A9 and A10
    - To make proposal equivalent, we need *another F bit* for BASE-R FEC ability
      - It makes implementation of BASE-R FEC effectively *optional*
    - For such purposes, there is a *simple workaround* without this flexibility
      - Just force link down, if the link is coming up in the BASE-R FEC mode
    - We cannot justify this non-standard usage to re-use existing 802.3bj design
      - Anyway, major logic change is required (i.e. 4 lanes → 1 lane x 4)
      - In the logic change, we should add BASE-R FEC that is mandatory for 802.3by
      - If BASE-R FEC is omitted, it is out of the standard, and we do not need to support it by Auto Negotiation (it can be still interoperable with 802.3by without AN)
- If we abandon this small freedom, we gain huge simplicity

Summary

FUJITSU

- 2 PMD types are not necessary
  - Use of 2 PMD types just increases confusion
- We lose almost nothing with unified logic-based resolution
  - We sacrifice small freedom of non-standard usage of A9 to reject CR-S (BASE-R FEC) and enforce CR (RS-FEC)
    - There is a simple workaround for such usage
      - Just detect it and force link down
    - We cannot justify this usage to re-use existing 802.3bj design
- We gain huge simplicity with single PMD
- Single PMD with optional RS-FEC and mandatory BASE-R FEC should be sufficient

FUJITSU

Thank you