

# 400GE Lane Configurations v.s. FEC Options

Zhongfeng Wang and Ali Ghiasi



IEEE 802.3 Plenary Meeting  
July 2013, Geneva, Switzerland

# Introduction

- This work provides preliminary analyses for possible FEC schemes to be considered by 400GE
  - The generic host FEC expected to be used for following PMDs: CDAUI-16, CDAUI-8, 400G-SR16
  - 400 GbE PMD based on 4 lanes of serial 100 Gb/s PMD may require PMD specific FEC due to high gain and complexity
  - 400 GbE backplane may require more complex signaling such as DMT and the generic FEC may not be enough
- At this early stage, we don't even have an specific PMD under consideration with numerous unknowns: total number of physical lanes, total number of PCS lanes, modulation format, etc
  - This analysis provide hypothetical tradeoffs between theoretical coding gain, overclocking rate, and processing latency
  - This analysis can be helpful in determining physical lane and/or logic lane configurations
  - This analysis can also help guide us if there is enough benefit to define a new FEC optimized for 400 GbE instead of reusing 802.3 BJ FEC

# Physical /Logical Configurations

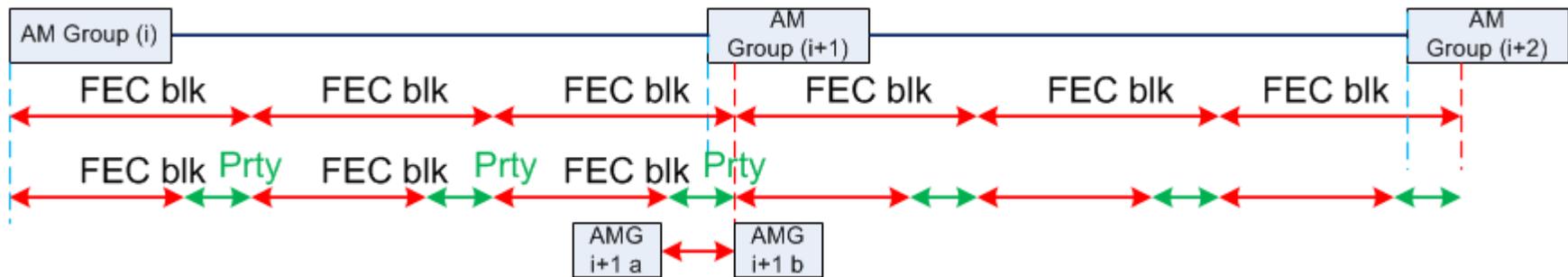
- For 400GE, based on current technology, 4 options may be considered for **the total number of physical lanes (PLs)**:
  - N=4
  - N=8
  - N=10
  - N=16
- Regarding **the total number of PCS lanes**, we may have following options:
  - L=4 (suit for 4 PLs)
  - L=8 (suit for 8 PLs and 4 PLs)
  - L=16 (suit for 4, 8 and 16 PLs)
  - L=20 (suit for 4 and 10 PLs)
  - L=24, or 48 (suit for 4, 8, and 16 PLs)
  - L=80 (suit for 4, 8, 10, and 16 PLs)

# Type of FEC Codes

- Based on current trend in the IEEE P802.3bj and IEEE P802.3bm, FEC will likely be included for 400GE
- Considering such a high speed requirement and general desire on low power and low latency, simple block codes such as BCH code or RS codes are promising candidate for FEC codes.
- Considering burst errors, RS FEC codes are well suited
  - BJ FEC is an RS FEC(528,514)
  - Early 400 GbE PMD implementation such as CDAUI-16 and SR-16 may not have error burst as the likely receiver will be based on CTLE but having a FEC with burst error is nice and will not limit future implementations.

# FEC Block Size v.s. PCS Lanes

- If encoding over multiple ( $L$ ) PCS lanes, multiple 66-b blocks are multiplexed into one data stream. Multiple AM blocks (i.e.,  $L$  AM blocks per AM group) are thus lumped together.



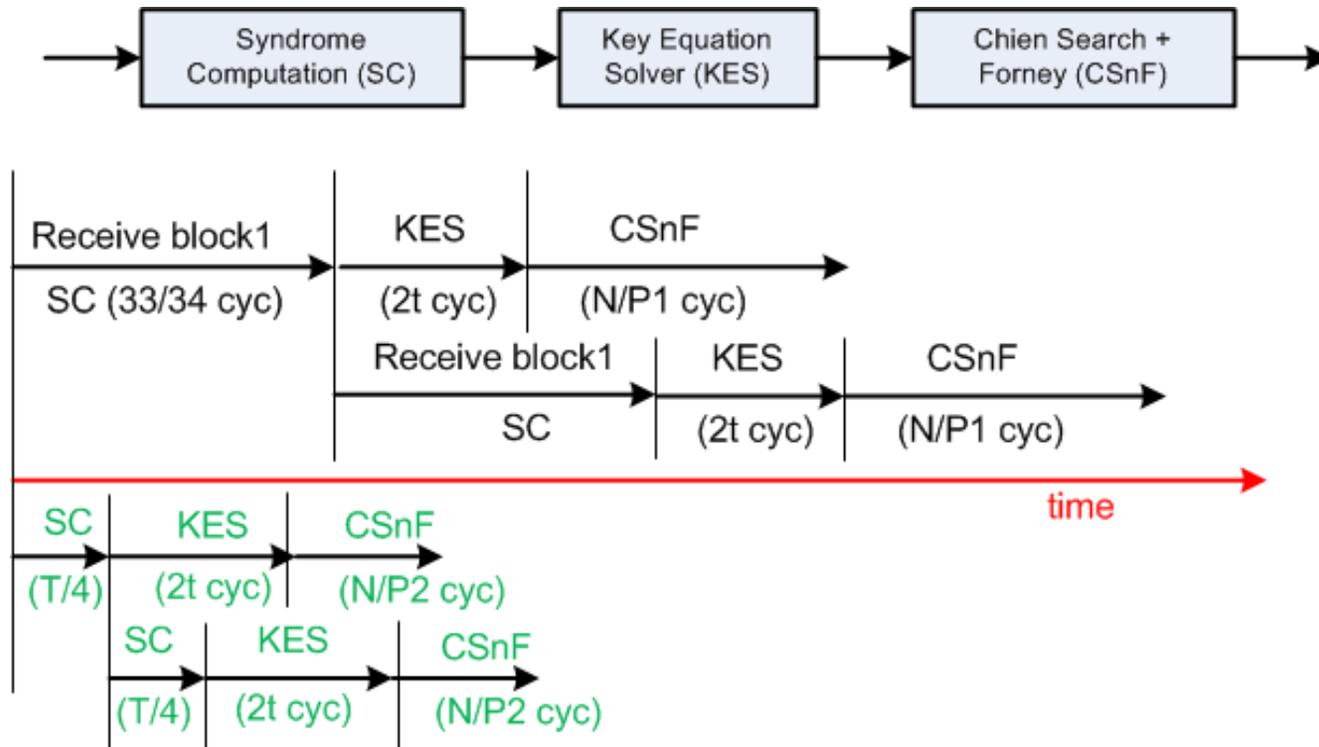
- It will cause implementation issue if the total number of bits between two AMGs is not multiple of FEC (source) block size **regardless using transcoding or not.**

# Alignment Marker (AM) Analysis

- Given a total of  $L$  PCS lanes, there're a total of  $L \times 16384 \times 66$  bits between two consecutive AM groups
  - Unless  $L$  is multiple of 5, FEC block size should not be a multiple of 10
  - Given other options of  $L$  ( $=4, 8, \text{ or } 16$ ), FEC block size should be a multiple of 4
- Considering RS codes defined over a finite field  $GF(2^m)$ :
  - For  $m=10$ , with no overclocking, RS( $t=7$ ) is the best option under certain constrains "BJ CL 91 FEC"
  - For  $m=11, m=12, m=13, \text{ or } m=14, \text{ or } m=15$ , RS code size [\*] will be a multiple of  $m$  (bits/symbol), which is not a factor of  $L \times 16384 \times 66$  when  $L=4, 8, 16, \text{ or } 20$  (considering either 256/257b or 512/513b transcoding)
  - For  $m=16$ ,  $L$  can be 4, 8, 16, 20, or 80
  - For  $m>16$ , overall latency and complexity will be a big concern.
- In brief, either  **$m=10$  or  $m=16$  is a good option.**
  - \* Adding dummy bits or shortening a code symbol is not considered here for ease of implementation.

# Decoding Latency for 400G FEC vs. 100G RS-FEC

- Parallel level in syndrome computation has to be linearly increased in 400G case in order to compute syndromes on-the-fly.
- Parallel level (**P2**) in Chien Search part should be increased in 400G case. But it may not be linearly increased (i.e., **4xP1**) considering implementation complexity.



# FEC Option I – Reuse bj FEC over 400G Data Rate

- To reuse bj FEC, it requires going with integer number of x10 pcs lanes
- For **m=10**, L=20 or 80 (PCS lanes)
  - RS(528, 514, t=7, m=10) (TC=256/257b ), same as 100G-KR4, 0% OC
    - NCG  $\approx$  5.7dB
    - Latency: transcoding + encoding + receiving block + decoding  $\approx$  45 ns
    - Reference: 100G-KR4 FEC gain  $\sim$  5.7 dB and latency : 85~95ns
  - RS(544, 514, t=15, m=10) (TC=256/257b ), same as 100G-KP4, 3% OC
    - NC  $\approx$  6.9dB
    - Latency:  $\approx$  70 ns
    - Reference: 100G-KR4 FEC latency: 95~105ns
- Reusing bj FEC across 400G PCS only reduces latency by about half since **decoding latency doesn't scale down as the block receiving time.**

# FEC Option II – Extended bj FEC over 400G Data Rate

- For **m=16**, L=4, 8, 16, 20, 80. Symbol size=16b and symbol interleaving is used for data distribution over multiple PLs.
  - Under no overclocking,
    - ❖ RS(528xK, 514xK, t=7xK), K=1, 2 or 4 (TC=256/257b )
      - RS(t=7), **similar to bj FEC except larger symbol size (16 vs. 10)**
        - NCG= 5.5dB
        - Latency: ~ = 50 ns
      - RS(t=14), double sized case
        - NCG ~ = 6.2dB
        - Latency: ~ = 98ns
      - RS(t=28), quadruple sized case
        - NCG ~ = 6.8dB
        - Latency: > 150ns
    - ❖ RS(528xK, 513xK, t=15xK/2), K=2, 4. (TC=512/513b )  
i.e., RS(t=15) and RS(t=30)
  - Under 3% overclocking (still ensure integer PLL)
    - ❖ RS(544xK, 514xK, t=15xK), K=1, 2 (TC=256/257b )
      - **RS(t=15), NCG ~ = 6.6dB, Latency ~ = 80ns**
      - **RS(t=30), NCG ~ = 7.4dB, Latency ~ = 100~ 160ns**
    - ❖ RS(544xK, 513xK, t=31xK/2), K=2 (TC=512/513b )  
i.e., RS(t=31), similar to RS(t=30) case.

# Summary of Coding Options

- The following options are provided for coding over 400G data rate.
- For **m=10**, **L=20 or 80**
  - Use 100G-KR4 FEC over 400G, OC=0%
    - NCG  $\approx$  5.7dB identical gain to BJ FEC
    - Latency:  $\approx$  45 ns, but latency was cut by  $\sim$  half
  - Use 100G-KP4 FEC over 400G, OC=3%
    - NC  $\approx$  6.9dB
    - Latency:  $\approx$  70 ns
- For **m=16**, L=4, 8, 16, 20, 80
  - Under no overclocking, RS(t=14),
    - NCG= 5.5dB
    - Latency:  $\approx$  50 ns
  - Under no overclocking, RS(t=14),
    - NCG  $\approx$  6.2dB
    - Latency:  $\approx$  98ns
- For **m=12**, L=24 (N=4, 6, 8) or **L=48** (N=4, 6, 8, or 16)
  - RS(528x2, 514x2, t=14), OC=0%
    - NCG  $\approx$  6.4dB (6.93dB for t=28)
    - Latency:  $\approx$  88ns

# Suggestions

- If using 16PCS, extended bj FEC should be considered
- If reusing bj FEC, 80 PCS lanes should be considered
- At this early stage not knowing all the upcoming PMD implementation, the PCS should not limit these future implementations
- The BJ FEC can address the need for generic host FEC, higher order modulation (HOM) expect to have an integrated high gain FEC
- The combination of FEC coding gain and/or latency is likely too little to redefine brand new FEC over 4 instantiations of BJ KR4 FEC.

## Future Work

- Power estimation will be provided in the next IEEE meeting
- Net coding gain over burst channels will be estimated and presented in next IEEE meeting.