Bypass Options for Concatenated FEC

Xiang He, Matt Brown

Huawei Technologies

Supporters

Cedric Lam, Google

Vishnu Balan, NVIDIA

Hao Ren, Huawei

Introduction

- Latency is critical to AI and ML applications.

- Power limitation is another critical issue for B400G in data center network.

- This contribution proposes options on how to improve these aspects for where optical PMD BER level is well controlled.

Background

- RS(544,514) has been adopted for 200G/lane AUIs (C2C and C2M).

  - See motions_3df_2211.pdf

- Concatenated code with RS(544,514) as the outer code is under discussion.

  - bliss_3df_01b_2211, farhood_3df_02b_2211 both proposed BCH/Hamming inner codes with RS outer code.

- Low latency and low power are two critical requirements for certain applications.

  - Latency is key for ML/AI with short fiber links, as in simms_3df_01_2210

  - Data center network generally is tight on power and latency, stone_b400g_01a_210301, lam_b400g_01a_210720

- The architecture of concatenated FEC enables lower latency and power than segmented because RS(544,514) is not terminated inside optical modules.

- This contribution discusses options to further lower latency and power for concatenated FEC.

Revisit: Latency of Inner Code Decoder

- The decoding latency of inner BCH/Hamming code itself is minimal.
    - Short BCH/Hamming decoding latency is as low as 1~10ns depending on algorithm (HD or SD).
    - 800 GbE as defined in P802.3df uses 4×212.5 Gb/s throughput RS(544,514) decoders, with a decoding latency of ~75ns.
        - See page #8 of he_3df_01_220517.
    - The latency of the inner code itself is 1.3% –12.8% of RS(544,514).

| FEC code			Operating rate	Latency[1], ns	Relative Area
Outer Code	Hard Decision RS	2-way RS(544,514)	850G	51.2	~4.00
		2-way RS(544,514)	212.5G	89.6	1.00 (Synthesized, 7nm)
Inner Code	Hard Decision BCH/Hamming	BCH(144,136)	225G	1.6	0.003
		eBCH(76,68)	~240G[3]	1.6	0.002
	Soft Decision BCH/Hamming (LRP = 6)[2]	BCH(144,136)	225G	9.6	0.17
		eBCH(76,68)	~240G[3]	9.6	0.11

1: Latency is evaluated based on 1.25 GHz clock frequency (0.8 ns per cycle).
2: Latency and/or area will go higher along with the performance if more LRP is selected.
3: Extra overhead is considered for single carrier 800Gb/s coherent transceivers.

he_3df_01a_220308.pdf

Interleaver for Concatenated FEC
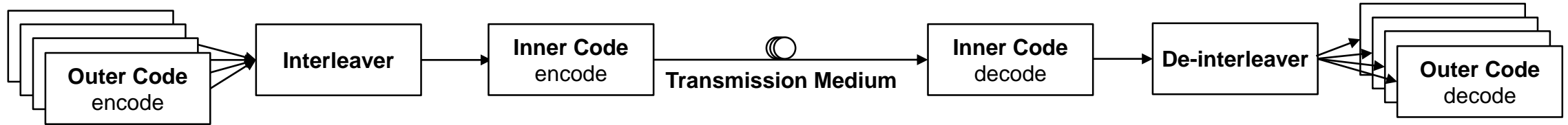
- Interleaver between outer and inner code can randomize the errors from inner code decoders.

- For block codes like RS FEC, convolutional interleaver is often used to lower latency.

  - Various convolutional interleavers have been discussed in the task force, with latencies varying from ~20ns to over 100ns, as in farhood_3df_02b_2211.

  - The interleaving depth of at least 12 RS codeword is recommended in multiple contributions.

Summary of SFEC (128,120) + Convolutional Interleaver : BER and Latency trade off for various operating modes

SFEC	Baud Rate	Convolutional Inter-leaver	Operating Mode	Encoder + decoder Latency	Pre-FEC BER
(128,120)	113.33Gbaud	High Latency mode	400G	~140ns	~4.8E-3
			800G ETC (2 way interleaved)	~140ns	~4.8e-3
			800G (4 way interleaved)	~ 56ns	~4.8E-3
			200G	~280ns	~4.8e-3
		Low Latency mode ** results in 0.25dB penalty in coding gain	400G	~56ns	~4.0E-3
			800G ETC (2 way interleaved)	~56ns	~4.0e-3
			800G (4 way interleaved)	~ 25ns	~4.0E-3
			200G	~110ns	~4.0e-3

# Tradeoff Between Latency and Pre-FEC BER Threshold

- Inner code itself does not require the convolutional interleaver to work.
  - Concatenated code performance without convolutional interleaver has been analyzed in [he_3df_01_2211.pdf](he_3df_01_2211.pdf).
- For links that has lower pre-FEC BER levels, convolutional interleaver can be bypassed.
  - 800 GbE PCS layer provides 4 codewords interleaving (likely for 1.6 TbE, too), which can provide moderate protection.
- For links that meets RS(544,514) threshold, the inner code can be bypassed completely.
- Tradeoff between latency and pre-FEC BER threshold can be made.
  - Bypass configuration can either be static or configurable.

BER Threshold	Bypass Convo. Interleaver	Bypass Inner Code	Inner Code Decoder: Soft or Hard Decision	Inner Code Total Latency
4.6E-3	No	No	Soft	50~300 ns
3.3E-3	Yes	No	Soft	5~10 ns*
6.1E-4	Yes	No	Hard	1~2 ns*
2.4E-4	Yes	Yes**	--	0 ns*

*Based on 200G/lane throughput, same for all Ethernet rates from 200 GbE to 1.6 TbE.
**Bypassing inner code will lead to different PMD rate.

High-level Block Diagrams, 800 GbE Example
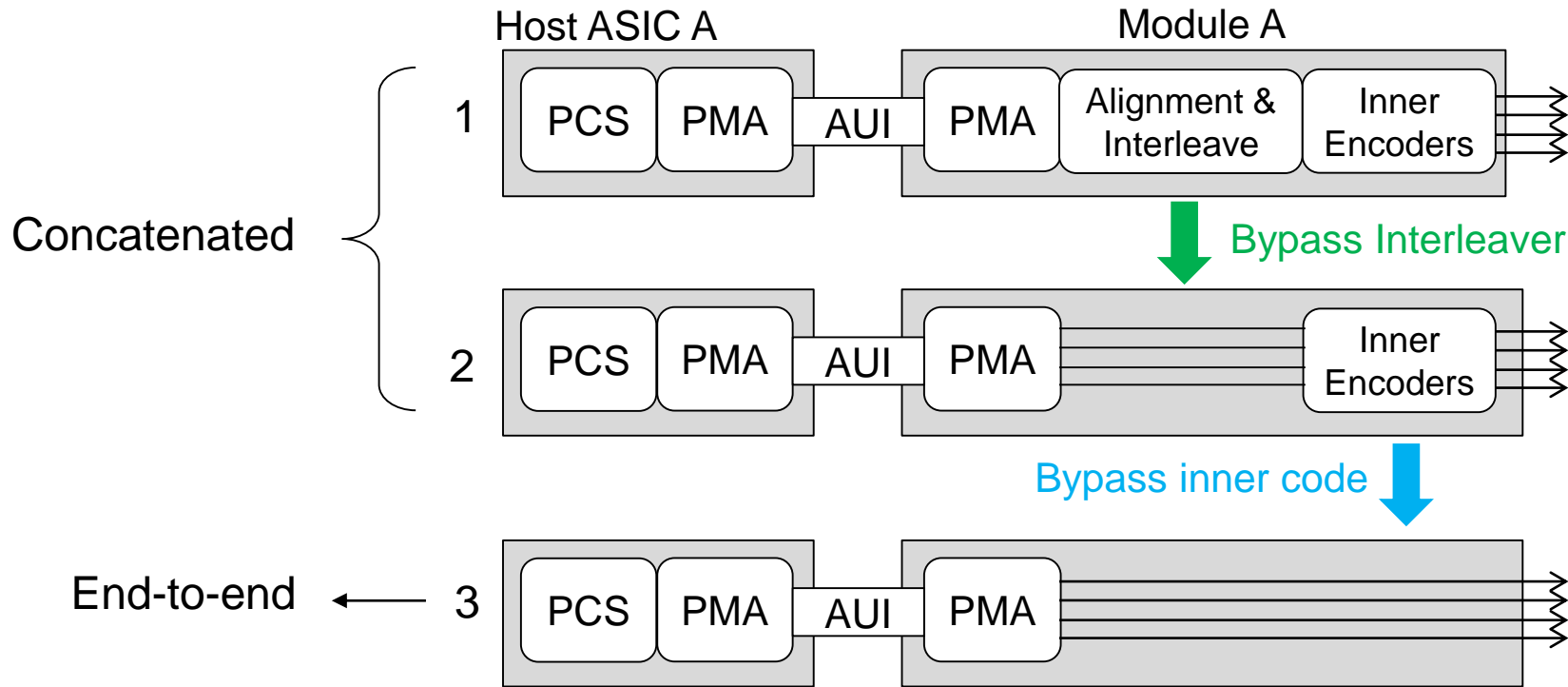
1. Both TX/RX modules may need to perform alignment and deskew depending on the interleaver design.
   - The alignment and deskew functions will require more logic and chip area (and power) inside module.
2. Interleaver bypassed. No alignment is required in both TX and RX modules. Lower latency and power.
   - Inner code can use self-sync to lock to the codeword boundaries.
3. Inner code bypassed. Essentially an End-to-End FEC.

Summary

- Concatenated code latency can be drastically reduced if the convolutional interleaver is bypassed.

  ▪ Convolutional interleaver is also a key contributor to power consumption for concatenated code.

- The standard should allow for bypassing of the convolutional interleaver.

  ▪ The concatenated code architecture should allow for interleaver bypass.

  ▪ Inner code may be bypassed when BER level is within RS(544,514) threshold.

  ▪ Methods to control the different bypassing options can be provided.

    – Static based on PMD or configurable.

Thank you