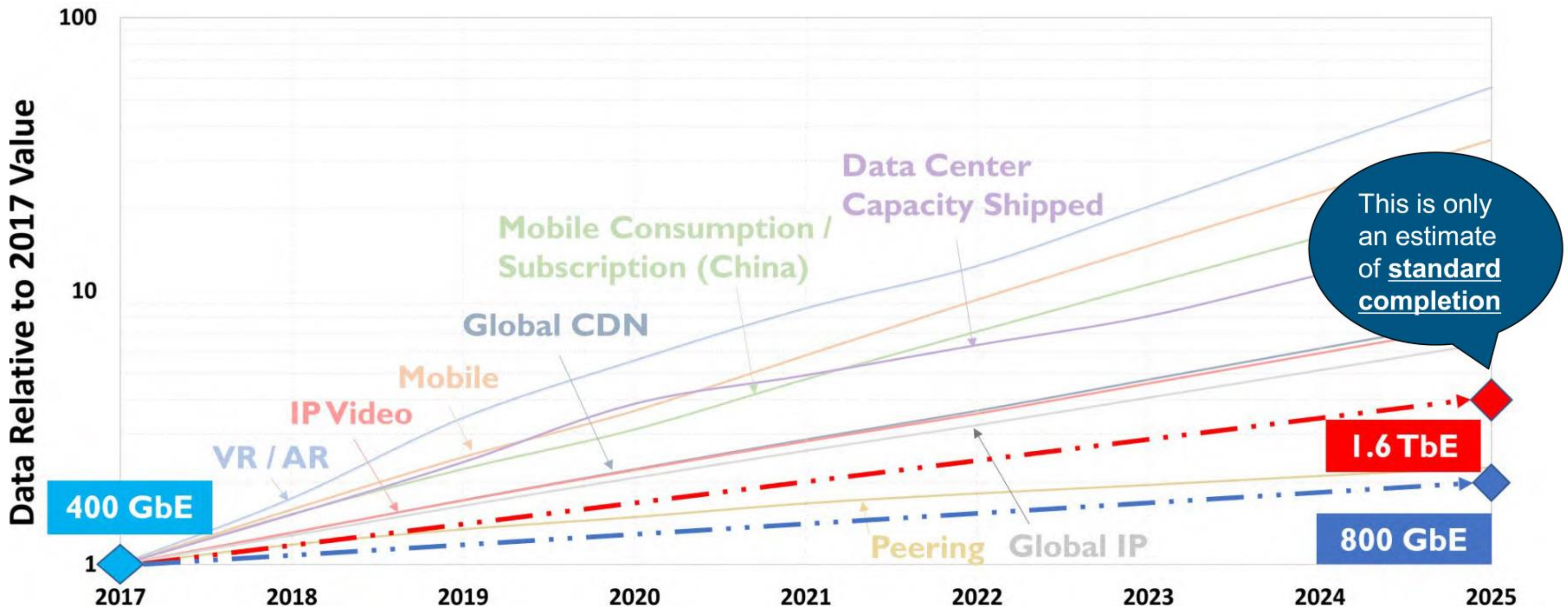
Technical Feasibility of Logic Layer to Support Rate Objective

Xinyuan Wang, Xiang He, Hao Ren

HUAWEI

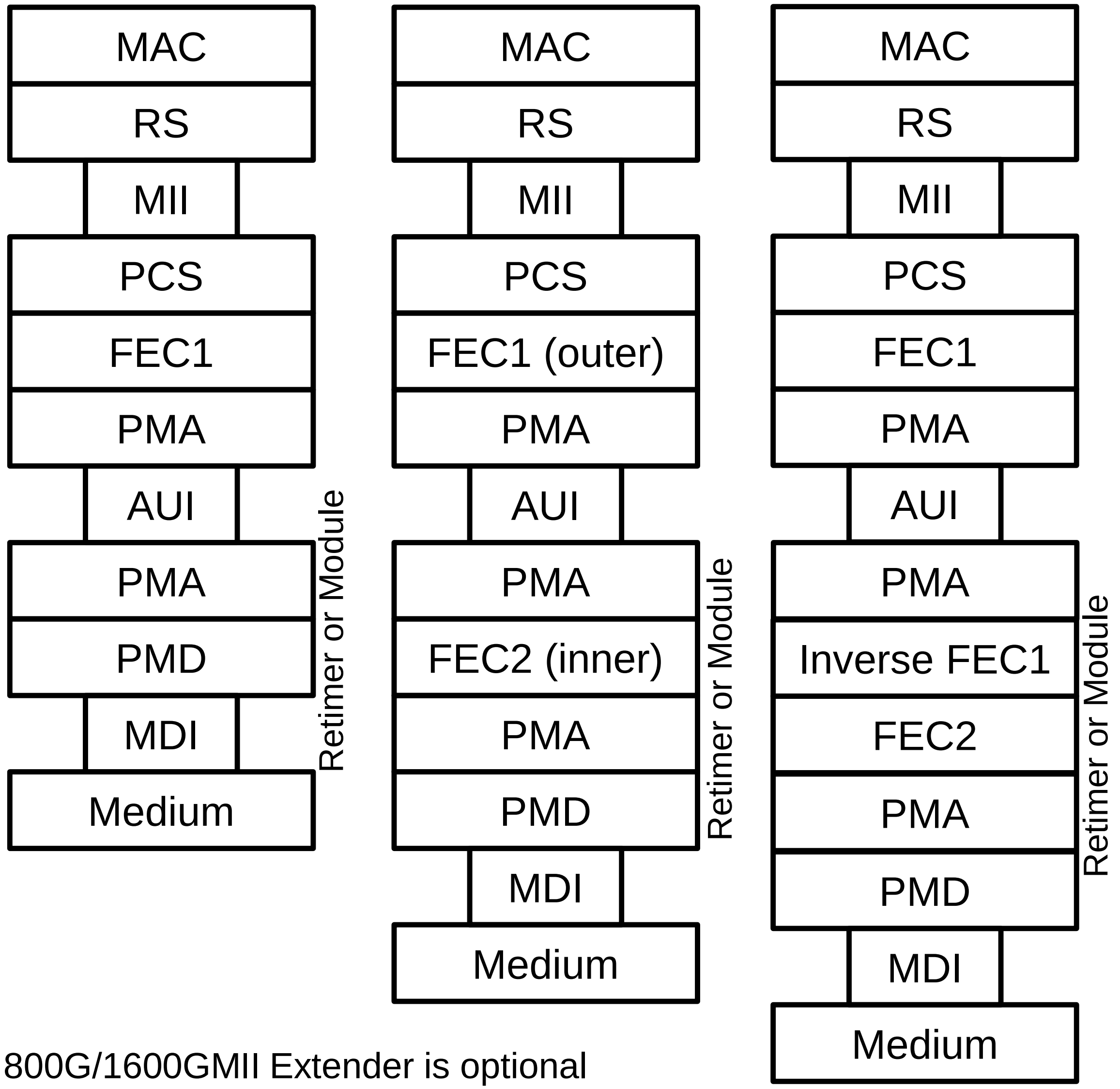
Background

- In Beyond 400 Gb/s Ethernet call for interest, general consensus was reached to start the debate for the next rate beyond 400 Gb/s.
  - 800GbE and/or 1.6TbE are potential candidates.



- In this contribution, we investigate MAC/PCS approach to support feasibility of potential 800GbE and 1.6TbE objective

Observation on 800GbE/1.6TbE Logic Architecture

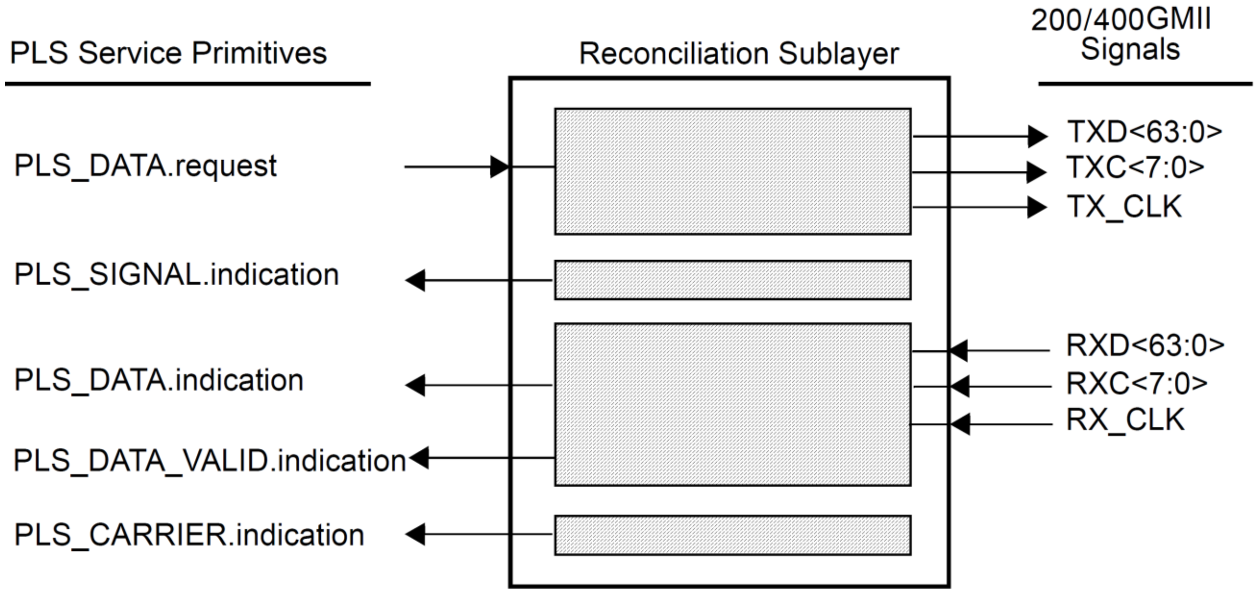


800G/1600GMII Extender is optional

- MAC: Similar as previous 200/400GbE
- RS/MII: CL81/117 with 64bit data and 8bit control
- PCS:
  - Encode/Decode: CL82/119 with 64B/66B
  - Scramble
- FEC:
  - Algorithm: RS FEC, Concatenative, Product, Convolution?
  - Architecture: End to End, Segment by Segment, Encapsulation
  - Implementation: Soft/Hard decision, Interleave, Parallelism, etc
- PMA: Bit mux or Block mux

RS and MII and 64B/66B Encode/Decode

- RS/MII: CL81/117 with 64-bit data and 8-bit control is reasonable. Further extending RS&MII data bus to larger than 64-bit, e.g. 128b/16Byte, is not doable, because it will violate Deficit Idle Counter mechanism and minimum IPG requirement of 12Byte, thus compromise line rate transmission in Ethernet.
- For 7nm node ASIC: 640bit@1.29GHz parallel implementation is achievable to enable 800Gb/s 64B/66B encode/decode.
- Forecasting 7nm and ≤5nm node ASIC: 1280bit@1.29GHz parallel implementation is feasibility to enable 1.6Tb/s 64B/66B encode/decode.



Reconciliation sublayer for 200/400GbE

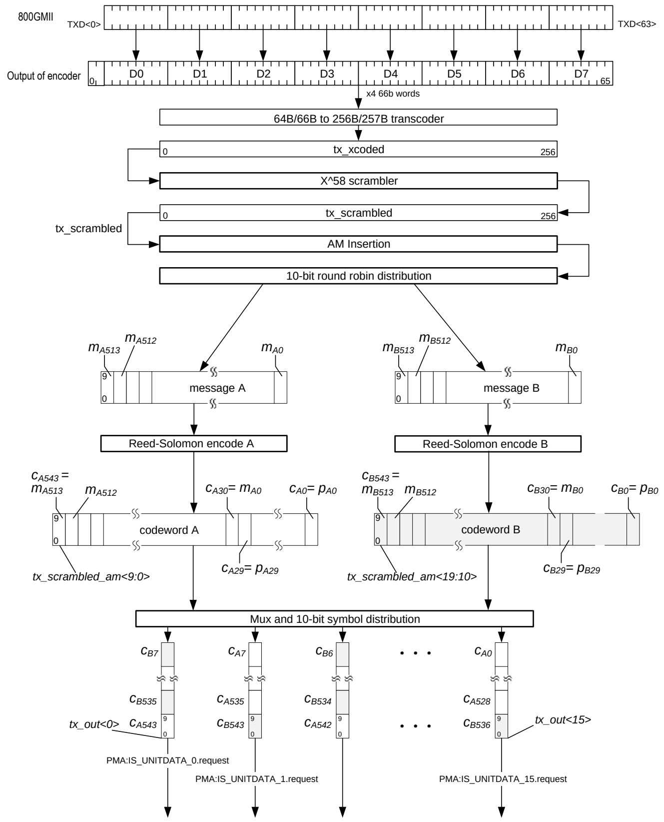
Assumption of Electrical/Optical Lane for FEC Analysis

FEC Approach #	AUI Interface	Optical Lanes	Reach
A End to End	800GbE: 8X100Gb/s, 4X200Gb/s 1.6TbE: 16X100Gb/s, 8X200Gb/s	800GbE: 8X100Gb/s 1.6TbE: 16X100Gb/s	50m-500m
B Encapsulation		800GbE: 4X200Gb/s 1.6TbE: 8X200Gb/s	50m-2km
C Segment by Segment		800GbE: 1X800Gb/s 1.6TbE: 2X800Gb/s, 1X1600Gb/s	10km+

Further explain on subsequent slides

FEC Approach A: End to End with RS(544,514)

- Assume refer to current specification of 802.3bs/cu/ck for 100G/s per lane, it is feasibility to double the rate of CL119 Architecture for 400GbE to achieve 800Gb/s capability of 7nm node ASIC, 640bit@1.33GHz for RS(544,514) decode with the following advantage:
  - Two 400Gb/s capability code words interleave to be fully backward compatible 802.3bs/cu/ck specification for 100G/s per lane, further lower power
  - Lower latency comparing to 400GbE with 12.8ns Versus 25.6ns for block time of RS(544,514) decode
  - 8 FEC Lanes, low complex and permit 100Gb/s and great per lanes AUI, Electrical/Optical Medium
  - For future ≤5nm node ASIC: same architecture with 1280bit@1.33GHz can enable 1.6Tb/s throughput RS(544,514) decode with 16 FEC Lanes



800GBASE-R Transmit bit ordering and distribution

FEC Approach B: Encapsulation with Concatenative Scheme

- As higher BER for 100Gb/s+ per lane is expected (e.g. 200Gb/s per lane), higher gain FEC comparing to ~6.4dB for RS(544,514) may be necessary. Concatenated FEC with >8dB is popular in industry.

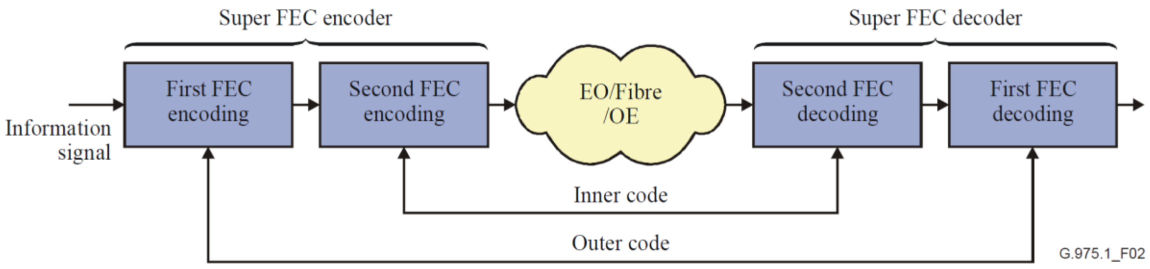


Figure 2/G.975.1 – Outer code and inner code

Table I.1/G.975.1 – Overview of super FEC schemes

Subclause	FEC scheme	
	Concatenated or non-concatenated	**Used FEC code**
I.2	Concatenated FEC	Outer code: RS(255,239) Inner code: CSOC (n_0/k_0 = 7/6, J = 8)
I.3	Concatenated FEC	Outer code: BCH(3860,3824) Inner code: BCH(2040,1930)
I.4	Concatenated FEC	Outer code: RS(1023,1007) Inner code: BCH(2047,1952)
I.5	Concatenated FEC (Soft Decision capable)	Outer code: RS(1901,1855) Inner code: Extended Hamming Product Code (512,502) × (510,500)
I.6	Non-concatenated FEC	LDPC Code
I.7	Concatenated FEC	Two orthogonally concatenated BCH codes
I.8	Non-concatenated FEC	RS(2720,2550)
I.9	Concatenated FEC	Two interleaved extended BCH(1020,988) codes

Refer to: ITU-T G.975.1(2004) Appendix I

- Encapsulation approach is compatible with Concatenative FEC scheme in 800GbE/1.6TbE era.

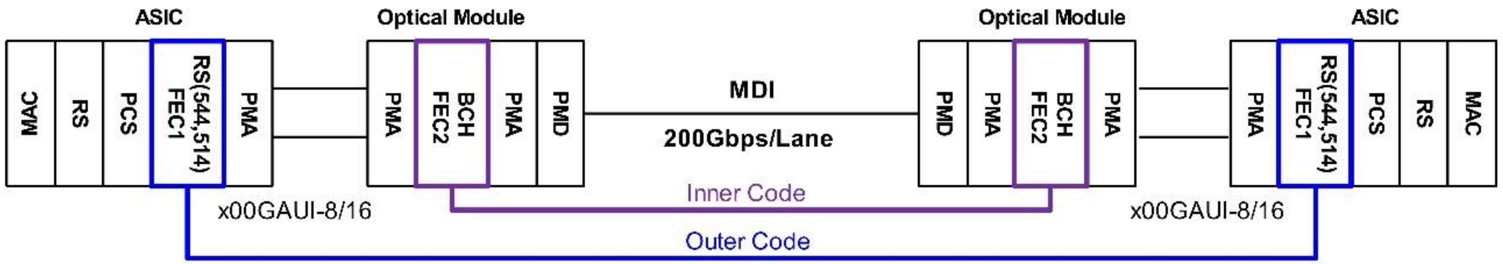
FEC Approach B: Encapsulation with Concatenative based on RS(544,514)

- Some examples for Concatenative scheme based on RS(544,514) as outer code, short code RS and BCH FEC as inner code.

Code: (n, k, t)	Galois Field (2^m)	clk @~600MHz Latency(ns)				clk @~1.25GHz Latency(ns)			
		HD	SD, 1x	SD, 2x	SD, 4x	HD	SD, 1x	SD, 2x	SD, 4x
BCH(126,119, 1)	2^7	3.2	30.4	17.6	11.2	1.6	15.2	8.8	5.6
BCH(144,136, 1)	2^8	3.2	30.4	17.6	11.2	1.6	15.2	8.8	5.6
BCH(180,170, 1)	2^(10/9/8)	3.2	30.4	17.6	11.2	1.6	15.2	8.8	5.6
BCH(360, 340, 2)	2^(10/9)	3.2	30.4	17.6	11.2	2.4	16	9.6	6.4
BCH(720, 680, 4)	2^10	4.8	32	19.2	12.8	4	17.6	11.2	8
RS(544,514, 15)	2^10	100.8				76			
RS(576,514, 31)	2^10	251.2				151.2			

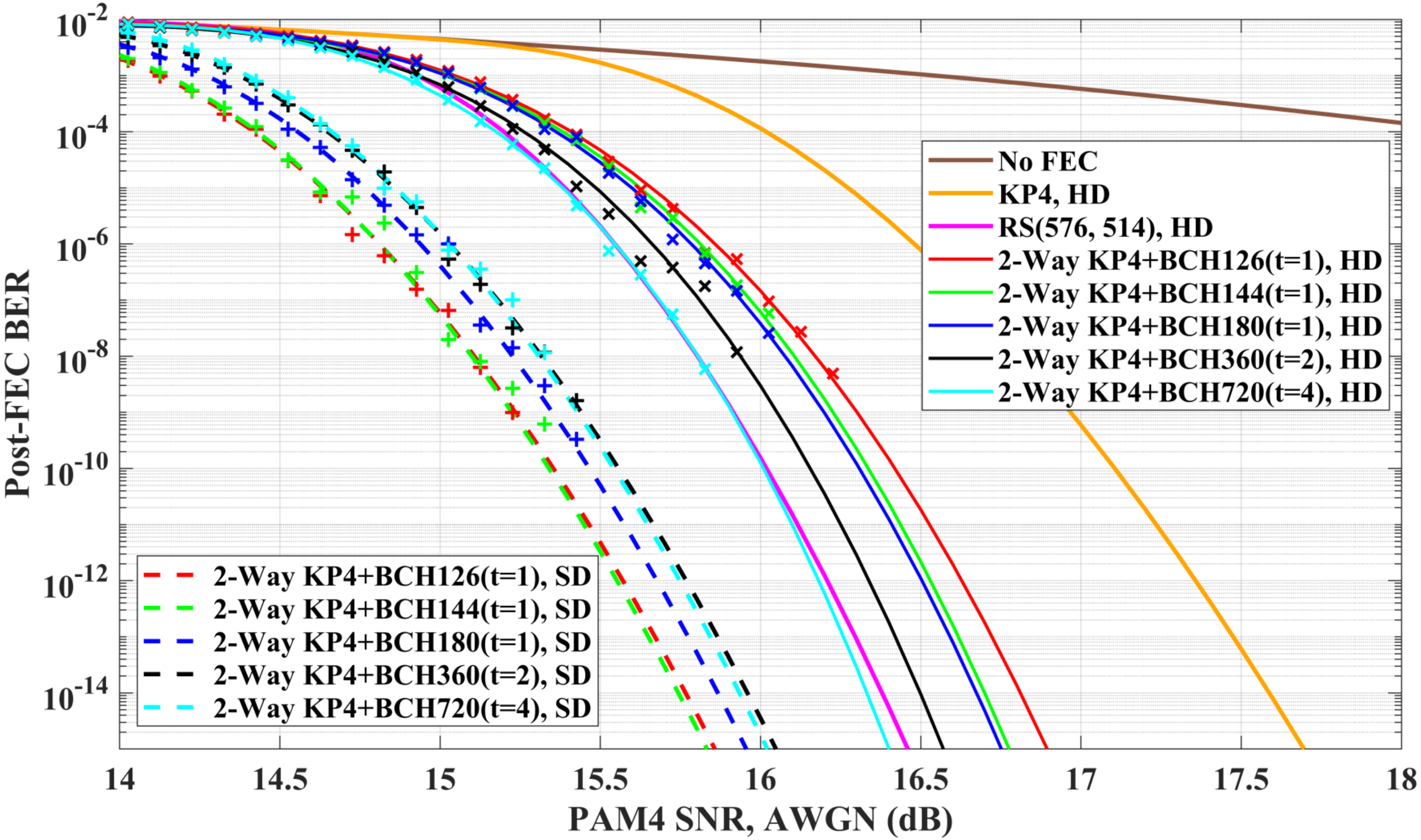
All Latencys are based on 200Gb/s through decoders

- Implementation scenario for Concatenative based on RS(544,514)



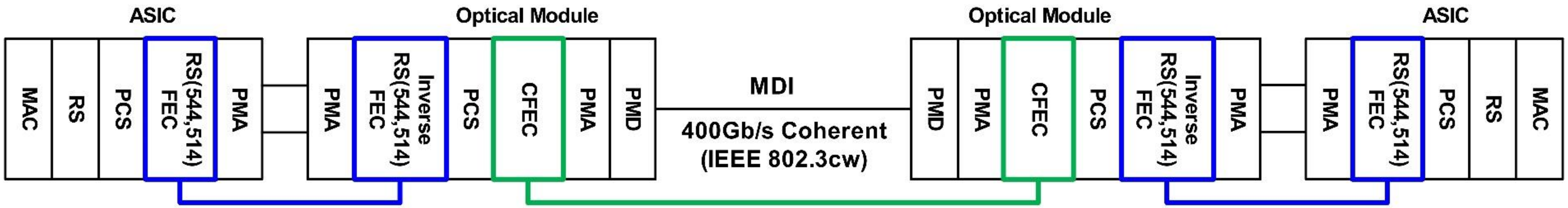
FEC Approach B: Encapsulation with Concatenative based on RS(544,514)

- RS(544,514) based concatenative BCH soft decision code approach can support both 100Gb/s and 200Gb/s per lane scenarios and interoperating.



FEC Approach C: Segment by Segment with Product or Convolution Code

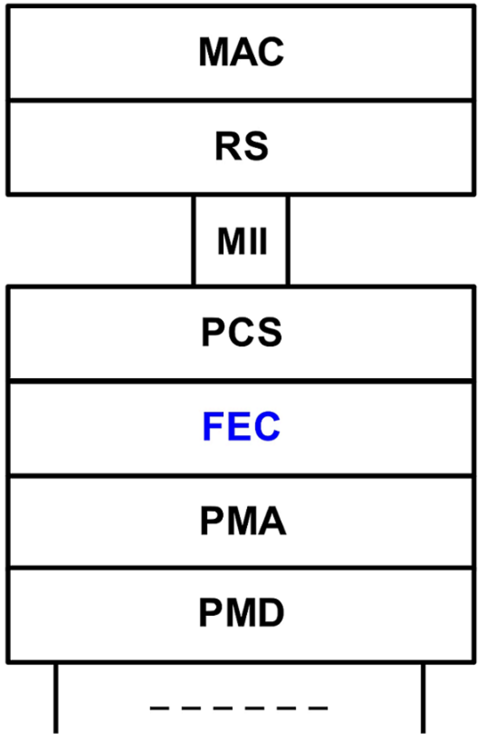
- For 10km+ reach PHY of 800G/1.6TbE, coherent is an approach and it will require much higher gain FEC with NCG ~10dB. Product code and convolutional code are already used in industry.
- The Segment by Segment FEC scheme in 802.3cw is a good example. RS(544,514) FEC covering the C2M interface is terminated inside the module, and a new FEC is added.



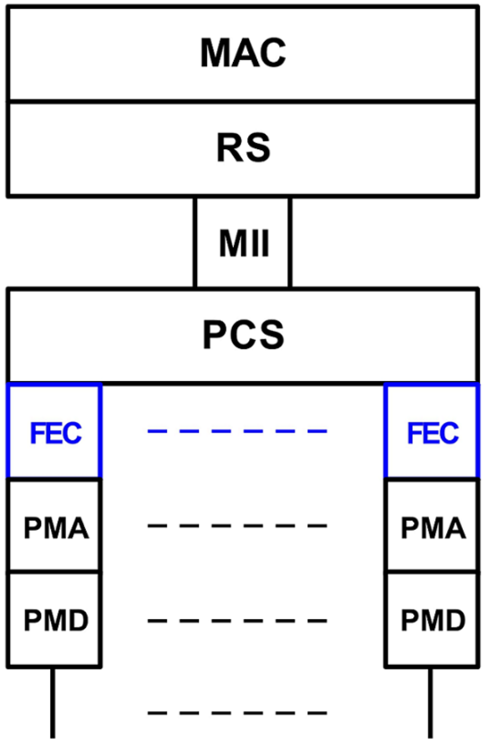
- Information for one of product code operating at 400Gbps throughout at 7nm node ASIC with 1024bit@400MHz
  - Latency: <8us; Power consumption: <2W;
- With 5nm or 3nm process node expected, it is technically feasible to achieve 800Gbps+ throughput for a ~10dB FEC.

Building Block of FEC Architecture: Per PHY or Physical Lane?

FEC per PHY



FEC per
Physical lane



- FEC per PHY has lower latency than per lane, with higher data throughput
- FEC per Physical lane gives higher capability on burst errors, with lower data throughput
- Tradeoff is needed between FEC capability, latency, implementation cost, etc. This should be further investigated in Task Force after the rate objective(s) is determined

Scramble: Clause 49

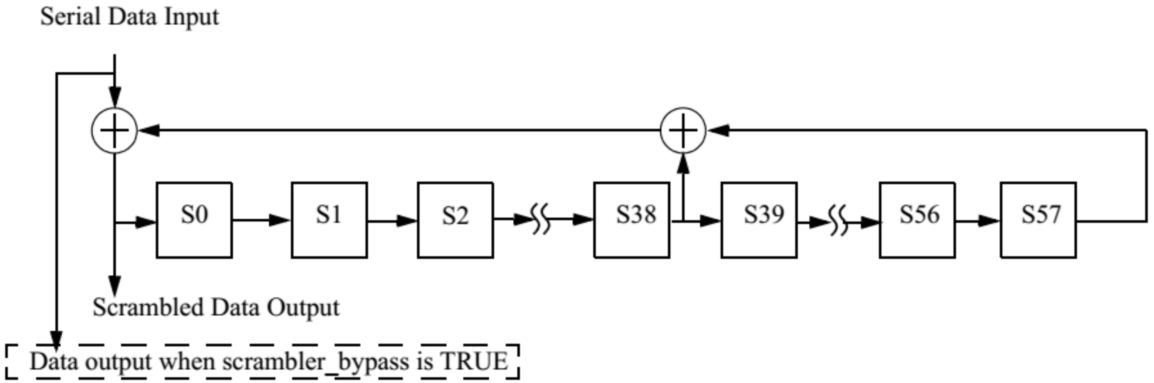
- Most possible to reuse scramble as in CL 49 for 802.3ae 10GbE, 802.3ba 40/100GbE, 802.3bs 200/400GbE, 802.3cd 50GbE

49.2.6 Scrambler

The payload of the block is scrambled with a self-synchronizing scrambler. The scrambler shall produce the same result as the implementation shown in Figure 49–8. This implements the scrambler polynomial:[8]

$$G(x) = 1 + x^{39} + x^{58} \quad (49\text{–}1)$$

There is no requirement on the initial value for the scrambler. The scrambler is run continuously on all payload bits. The sync header bits bypass the scrambler.



NOTE—Scrambler_bypass is only required to support EEE capability.

Figure 49–8—Scrambler

PMA: Bit Mux or Block Mux

- Bit Mux is preferred
  - Enable protocol agnostic optical module and friendly reuse in non-Ethernet interconnect area
- Block Mux:
  - Protocol aware optical module as delimiter block boundary
  - Some better performance in FEC for burst error than Bit Mux

Further Work Related with Logic Layer Architecture

- **PHY BER?**
  - 1E-14 comparing to 200G/400GbE with 1E-13, 40/100GbE with 1E-12?
- **200Gb/s per lane AUI Interface BER, 1E-5?**
  - 1E-5 for 50Gb/s and 100Gb/s per lane in Annex 120E.1.1/120G.1.1
- **FEC Approach A/B/C with long term evolution**
  - PMDs solution and operate over AUI interface simultaneously
- **FEC Lane number, SerDes Rate?**
  - 16X50Gb/s Versus 8X100Gb/s for 800GbE?
  - 16X100Gb/s Versus 8X200Gb/s for 1.6TbE?
- **Lower latency and power consumption by advanced process technology**
- **Breakout**

Summary

- From logic layer technical feasibility perspective, support the following potential objective:
  - Support a MAC data rate of 800 Gb/s
  - Support a MAC data rate of 1600 Gb/s
- Various FEC architectures can be used to achieve the data rate and coding gain necessary for higher Ethernet rates

Acknowledge

The authors would like to thank Professor Zhongfeng Wang and Jun Lin from Nanjing University for supporting FEC research work to this contribution

The authors would like to thank Matt Brown from Huawei Canada for reviewing and contributing to this contribution

Thank you