Some Consideration of Stronger FEC in 400GbE

Xinyuan Wang, Yu Xu, Wenbin Yang

HUAWEI TECHNOLOGIES CO., LTD.

IEEE 802.3bs 400 GbE Task Force

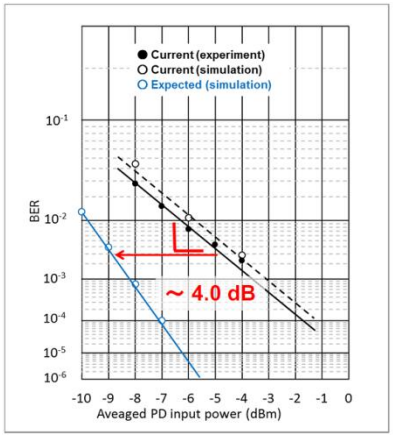
HUAWEI

Introduction and background

- In this presentation, we investigate stronger FEC schemes, their limitations and how to adapt them in the logic layer architecture.
- In many presentations the reuse of 802.3bj RS FEC (KR4 and KP4) was assumed with extension to 4X parallelism to support 400GbE packet flow;
- Many presentations on SMF PMDs mention the use of significantly stronger FEC (e.g.BCH) to improve available loss budget and satisfy system MTTFPA/FLR.

10km	8x50G NRZ	8x50G PAM4	8x50G PAM4	4x100G PAM4	4x100G DMT-k
Source	Mod	Mod	DML	Mod	25G DML
Grid	LAN WDM	LAN WDM	LAN WDM	LAN WDM	LAN WDM
FEC	BCH	BCH	BCH	BCH	BCH
Operating BER	1.0E-03	1.0E-03	1.0E-03	1.0E-03	1.0E-03
TX analog BW *	21	21	21	21	21
RX analog BW *	42	21	21	32	21
RX FFE taps or FFT pts.	3	3	3	13	k
TDP dB	**2.3**	**0.7**	**1.5**	**2.5**	**3.0**

cole_3bs_02b_0914



116 Gbps
SC: 256
Target BER 3.3E-3
BCH(9193,8192)

takahara_3bs_01a_0914

Benefits of Stronger FEC in 400GbE PMD?

- Example: Power limited system, 4X100Gbps PAM4 for 2km SMF

Realistic loss budget (Black & White analysis) for 4x100G PAM4 configuration

	Realistic specification for 2km duplex SMF	Realistic specification for 500m PSM4 SMF	Unit
Tx OMA (01-00) min Specification Value	-6	-4	dBm
TDP	1	1	dB
Tx OMA (01-00) – TDP min	-7	-5	dBm
Channel insertion loss Specification Value	4	4	dB
Rx ROP OMA (01-00) with KP4 FEC Specification Value	-6	-8	dBm
Available channel loss	-1	3	dB

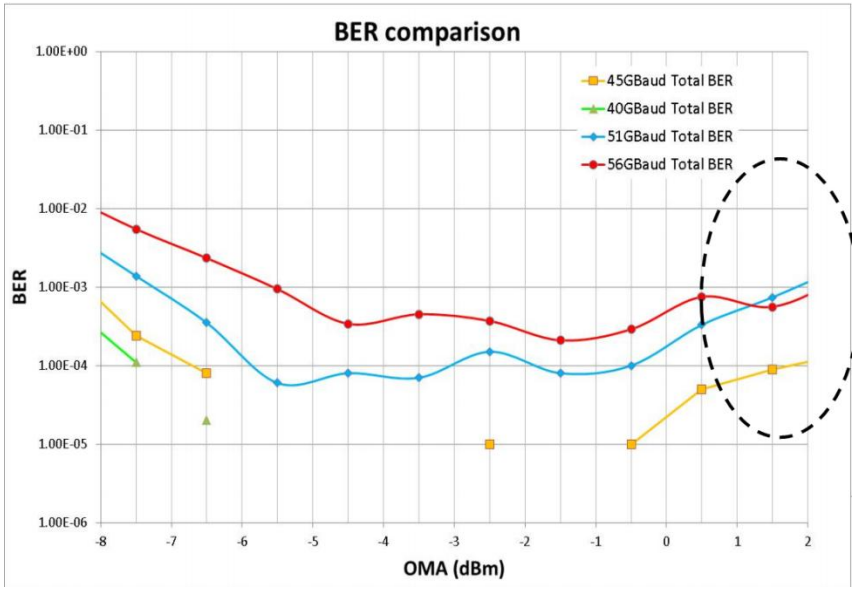
For 2km duplex SMF the "gap" in this budget seems too big to be bridged.
If reconfirmed then 4x100G PAM4 may only be useable for 500m PSM4.

stassar_01a_0914_smf

- For 4x100Gbps PAM4 / 2km need significantly stronger FEC with 10 dB more coding gain (compared to 6.6 dB for KP4 FEC).
- For 4x100Gbps PAM4 / 500m need FEC with 2 dB more coding gain

Is stronger FEC in 400GbE PMD the magical solution?

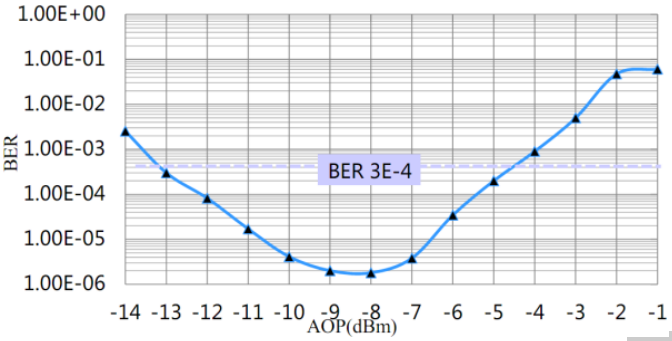
- If raw BER floor is present in SNR limited systems and the FEC operating point is not too far above the floor, will this floor stay at this level when system parameters vary (jitter, voltage, temperature, dispersion, pattern) or will it strongly move up/down?
- For multi-vendor interoperability, essential BER floor should be “safe distance from FEC operating point.



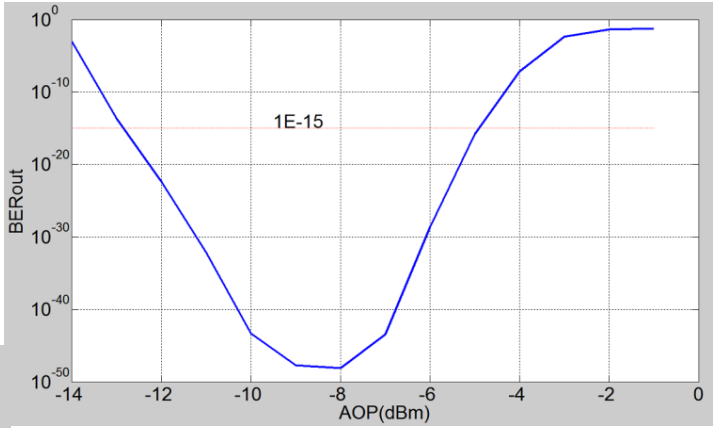
mazzini_3bs_01_0914

What is maximum acceptable BER Floor?

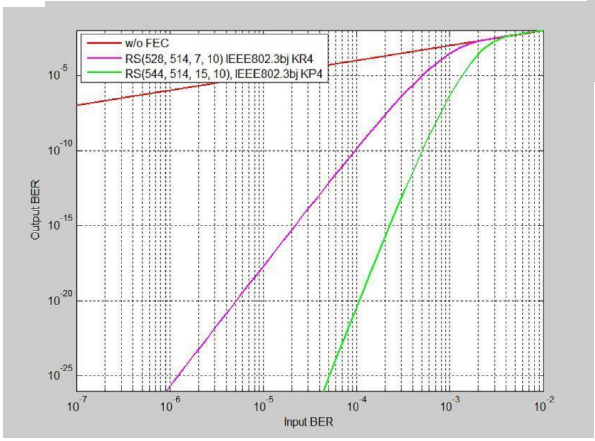
- It should depend on the required BERpost from MTTFPA objective.
- Taking RS(544, 514) as an example, generally a BER floor relatively close to the pre-FEC BER curve, will move to a significantly lower point after applying FEC.



After KP4 FEC



xu_3bs_01_0714



What will impact (stronger) FEC capability in 400GbE PMDs?

- Error Behaviour: Random vs Burst Error?
  - "Error Distribution in Optical Links" from "anslow_01_1107" in 802.3ba.

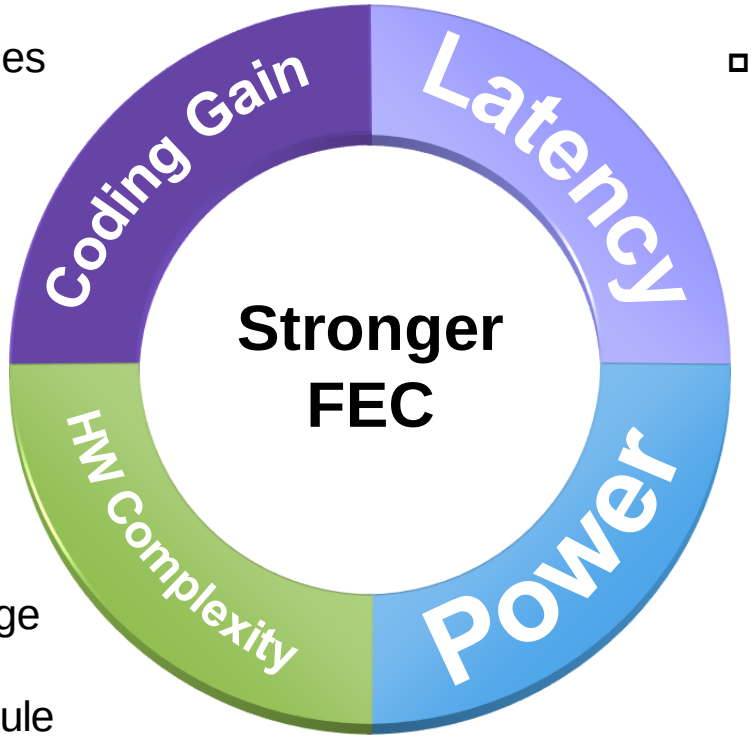
    It is expected that error distribution statistics will be approximately random for optical links where the BER is dominated by:
    - Noise
    - Chromatic dispersion
    - Modest Self Phase Modulation (SPM)
    - Polarisation Mode Dispersion (PMD)
    - All except extreme amounts of Four Wave Mixing (FWM)
    - If the BER is limited by some form of interference (e.g. spikes getting through the power supply filtering) then this could cause bursts of errors. This is very hard to quantify.
    - If EDC were part of the solution, Decision Feedback Equalisation (DFE) or MLSE may cause error bursts. (See next slides for some MLSE based EDC results).
- High order modulation schemes (PAM4, DMT) are proposed in 802.3bs. What is the behaviour of error distribution? Or, can we refer to error propagation model in electrical link methodology with Gilbert burst error model?

General requirements for stronger FEC

- Net Coding Gain(NCG):
  - Generic KR4/KP4 FEC with refer to gustlin_01_0112 in 802.3bj@1E-12 with <7dB coding gain. Stronger FEC in 400GbE will require >8dB NCG in general;
- Latency:
  - FEC coding/decoding latency is critical in some 400GE application scenarios. Latency is resulting from coding/decoding process, buffering for interleaving and de-interleaving process. Longer code lengths and deeper interleave/iteration will increase latency. How much latency is acceptable for datacenter applications?
- Overhead:
  - More overhead used for parity-check will get more NCG generally. If stronger FEC is integrated in Host ASIC, high over-rate of SerDes will impact the specification of chip-module interface. For example, CDAUI-16.
- Hardware Complexity:
  - Main portion of FEC implementation is decoder. Decoding complexity should be considered for embedding in Host ASIC/FPGA or silicon chip in module for Multiplexing/re-timer. Silicon implementation is not for free when comparing to optical technology cost, especially when integrated with other functions in one-chip.
- Power:
  - One of 400GbE main applications is in high density port line card/chassis. Power consumption and thermal dissipation will be a huge issue.

400GbE Stronger FEC tradeoffs

- Overhead Vs SerDes rate & technology feasibility.



- Latency in sensitive applications, such as Finance, DC,……. Especially for short reach solutions,100/500m.

- Difficult to be integrated in host ASIC or FPGA if large resource required.
- QSFP+ & CFP module with silicon chip embedded?

- Impact on small form factor module objective.

Stronger FEC example:

- FEC core size from:"langhammer_01_1014_logic",

Type	Codeword	Area (6LUT)	Relative Area
RS KR4	(528,514,7)	10654	1
RS KP4	(544,514,15)	26554	2.5
BCH[1]	(2858,2570,24)	106806	10
BCH[2]	(9193,8192,71)	425000	40

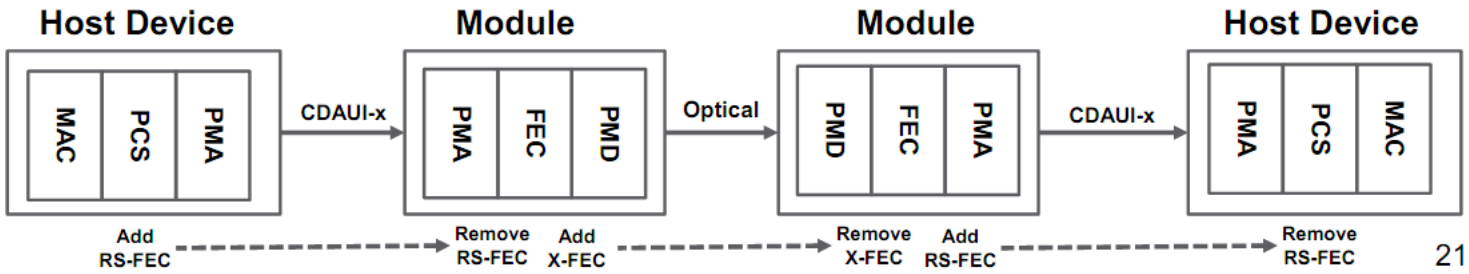
- Refer to modelling method in Langhammer's slide, further compare different RS FEC:

	CG@1E-13	NCG@1E-13	Over Clock	Relative Area
RS(528,514,7,10)	5.3942	5.2775	0	1
RS(544,514,15,10)	6.6357	6.3894	3.03%	2.9
RS(560,514,23,10)	7.3012	6.9289	6.06%	5.9
RS(576,514,31,10)	7.759	7.2645	9.09%	9.9
RS(592,514,39,10)	8.1076	7.494	12.12%	15

- BCH1 can tolerate 288bit random/correlation errors. BCH2 can tolerate 994bit random/correlation errors.
- Based on our evaluation, 400GbE MRC/PCS implementation (CRC32+16 Lanes PCS) will require ~200K LUT resource. If 4xKR4 FEC for 400GbE, the Ratio of "KR4 FEC: MAC/PCS" = 4:20.
- BCH FEC will consume most logic resource for monolithic FPGA.
- RS(560,514) or RS(576,514) is reasonable considering NCG benefit versus over-clock expense. Due to HW complexity and power, it still difficult to be integrated in Host side ASIC.

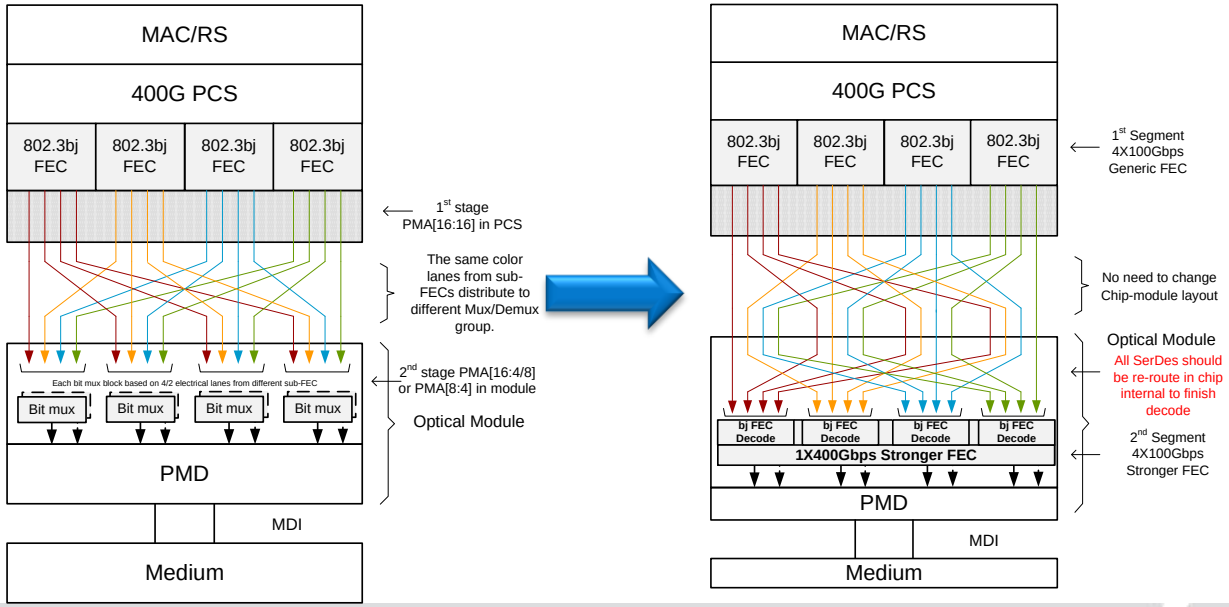
How to adapt Stronger FEC in 400GbE logic layer

- Based on “Segment by Segment” FEC strategy:



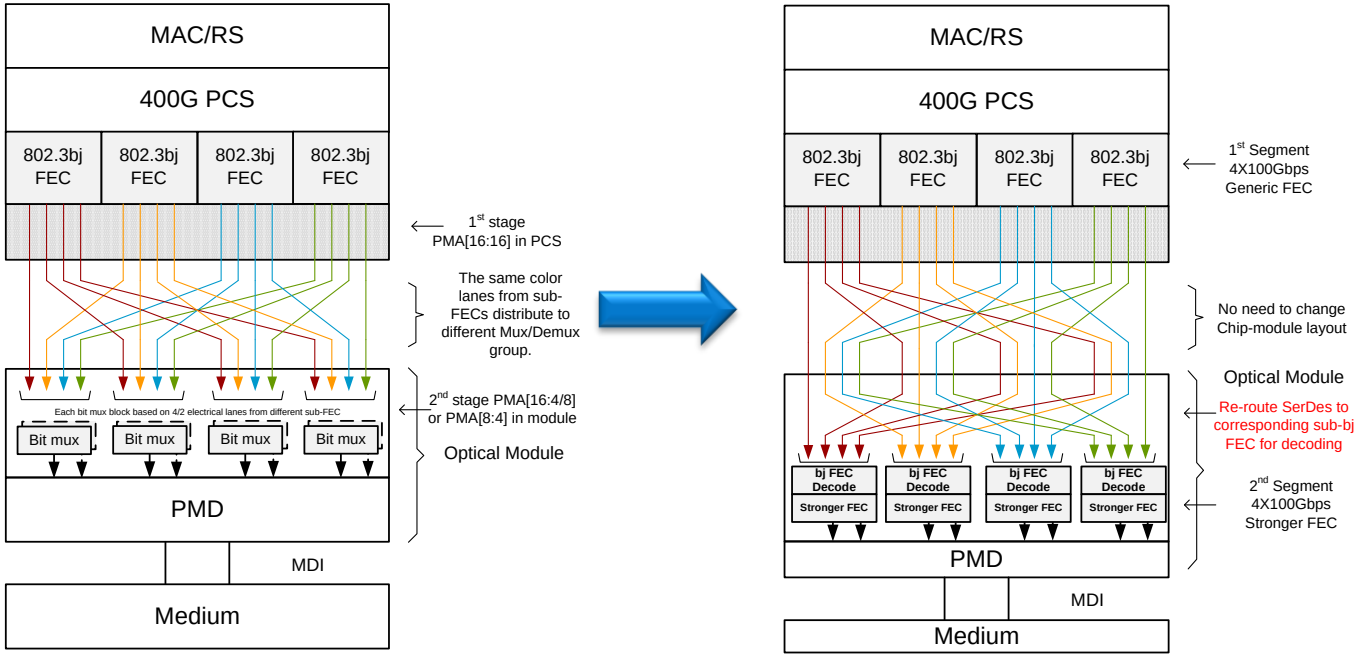
anslow_3bs_02_0914

- Scenarios 1: 1x400Gbps stronger FEC is implementation for optical module:



How to adapt Stronger FEC in 400GbE logic layer

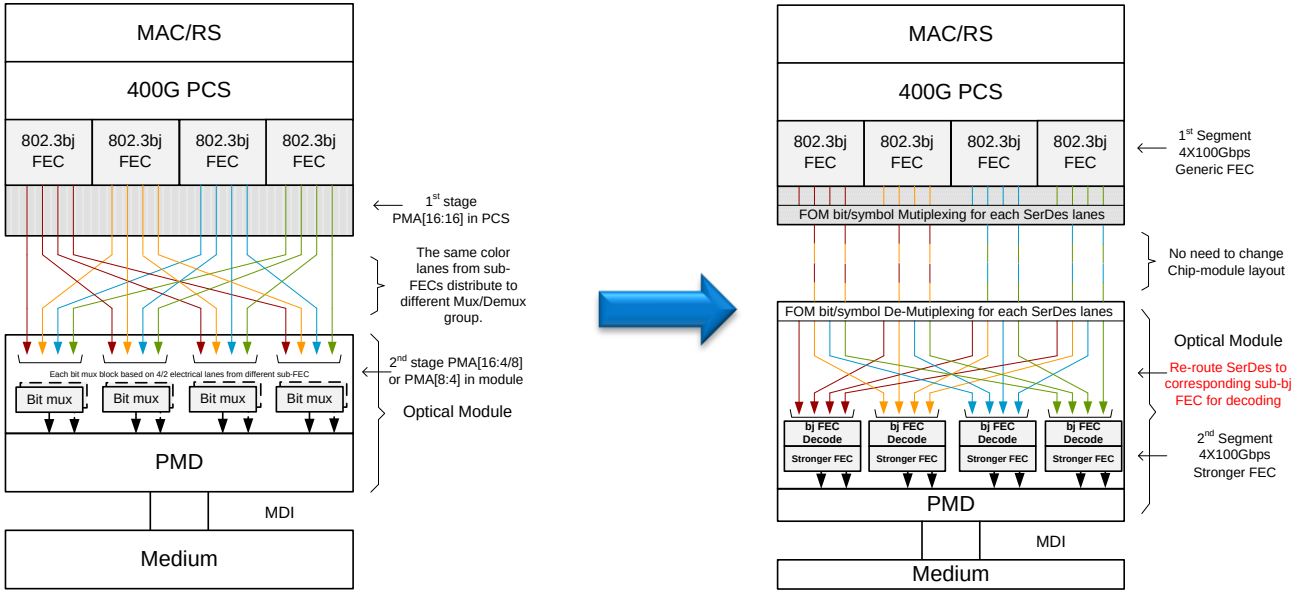
- Scenarios 2: 4x100Gbps stronger FEC implementation for optical module with CDAUI-16:



- Implementing 1x400Gbps stronger FEC by 4x100Gbps parallelism increases technical & economical feasibility and enables reuse of 100Gbps per lanes.
- Each 100Gbps stronger FEC connecting to one bj FEC respectively is more reasonable.

How to adapt Stronger FEC in 400GbE logic layer

- Scenarios 3: 4x100Gbps stronger FEC implementation for Optical module with CDAUI-8:



- For CDAUI-8/4 chip-module interface, the checksum of stronger FEC will replace original checksum bit of BJ FEC. This will require to identify each FEC lane from SerDes and re-route to respective BJ FEC decode block in silicon of module.
  - Non-FOM bit/symbol multiplexing is more easy to cooperate with stronger FEC, most like to support breakout. This will waste the FOM error enhance capability and lead to two different PMA architecture in end-to-end and segment-by-segment FEC strategy respectively.
  - Re-define Host FEC architecture, 16x25Gbps sub-bj FEC for unlimited FOM architecture? Can we endure the additional cost of latency and hardware complexity in this proposal?

Summary

- FEC, stronger than BJ FEC, will make 400GbE solutions significantly more complex.
- we should be careful to say “stronger FEC” is ready to be adopted. More detailed work needed in the future.
  - Stronger FEC embedded in Host ASIC is not cost effective as large logic resource.
  - If stronger FEC is embedded in silicon chip of optical module, it will limit options to get a small form factor and still be a big challenge in power&complexity.
  - Additionally stronger FEC on host board also makes line card designs more complex, difficult to implement.
- In the case of a stronger FEC only for the optical section, the scenarios to support this in the host PCS/PMA architecture will need to be investigated.

Thank you