MTTFPA for Uncorrected Transcoded FEC Blocks

IEEE 802.3bj interim meeting

May 17-18, 2012, Minneapolis, MN

Roy Cideciyan - IBM

Introduction

- For correlated errors at the decision-feedback equalizer (DFE) output, transcoded and forward error correction (FEC) encoded blocks have a sufficiently high mean time to false packet acceptance (MTTFPA) when FEC decoder is on.

- If transcoded and FEC encoded blocks have a sufficiently high MTTFPA when FEC decoder is off, we could always transcode and encode data using FEC and use FEC decoder only if necessary.

- The mode of always transcoding and FEC encoding data as an alternative to auto-negotiation based on cable loss is of special interest to copper cable channels for achieving 5m objective: *gustlin_01_0312.pdf*

- 256b/257b transcoding (TC): *cideciyan_01a_0312.pdf*

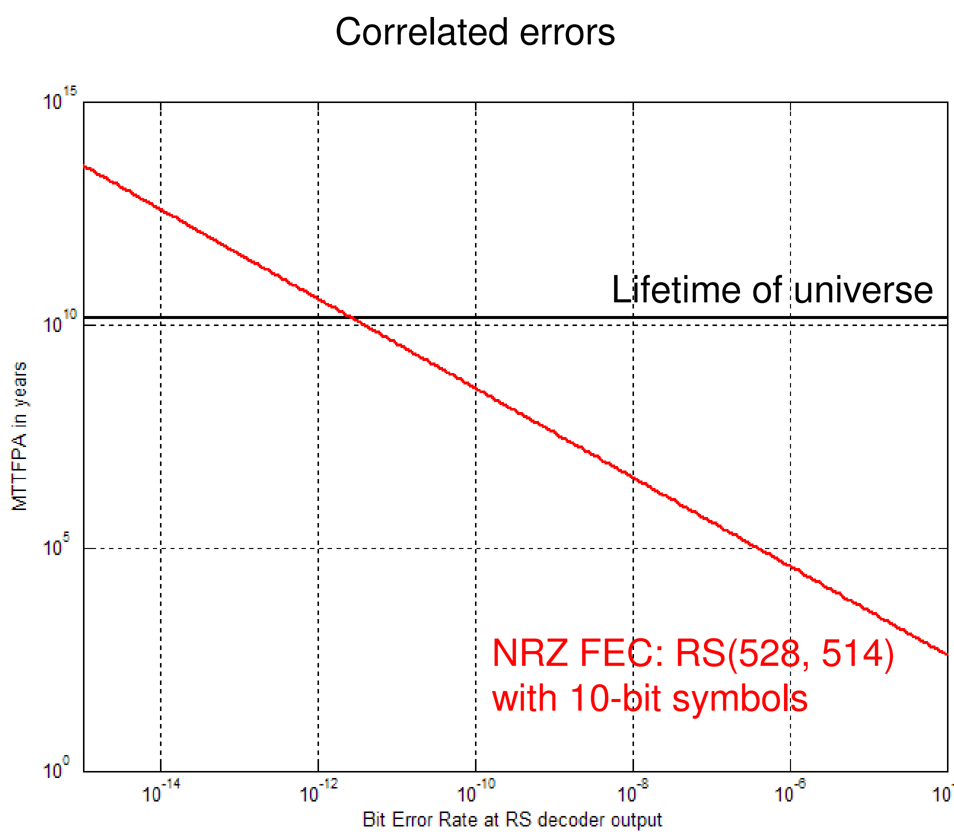
- MTTFPA for uncorrected transcoded FEC blocks is estimated

MTTFPA - FEC Decoder is On

- For correlated errors prob. that burst error continues is a=0.5
- MTTFPA for NRZ FEC RS(528,514) at BEI < 1e-12

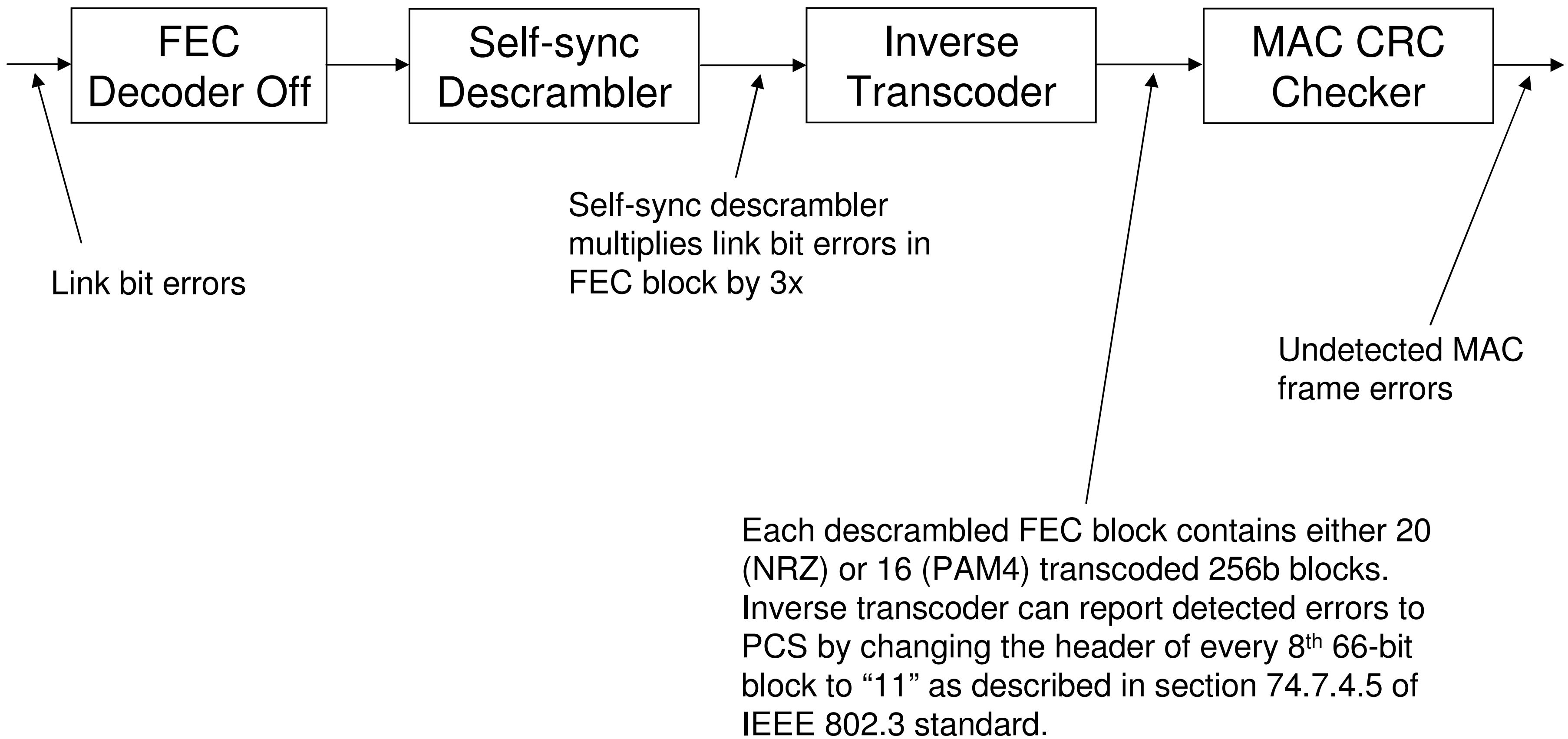
  ➔ Very good
- MTTFPA for NRZ FEC at BER ~ 1e-6

  ➔ Satisfactory
- hi_ber asserted true if BER > 1e-4 (clause 82.2.18.2.2)
- RS(444,412) for PAM4 has higher error correction capability: *brown_01a_0312.pdf*
- PAM4 FEC has better MTTFPA than NRZ FEC



cideciyan_01_0112.pdf

Undetected Frame Errors - FEC Decoder is Off



CRC Error Detection Capability

- IEEE standard CRC code has Hamming distance 4 for frame sizes used in ethernet and therefore can detect up to 3 bit errors in transmitted ethernet frames

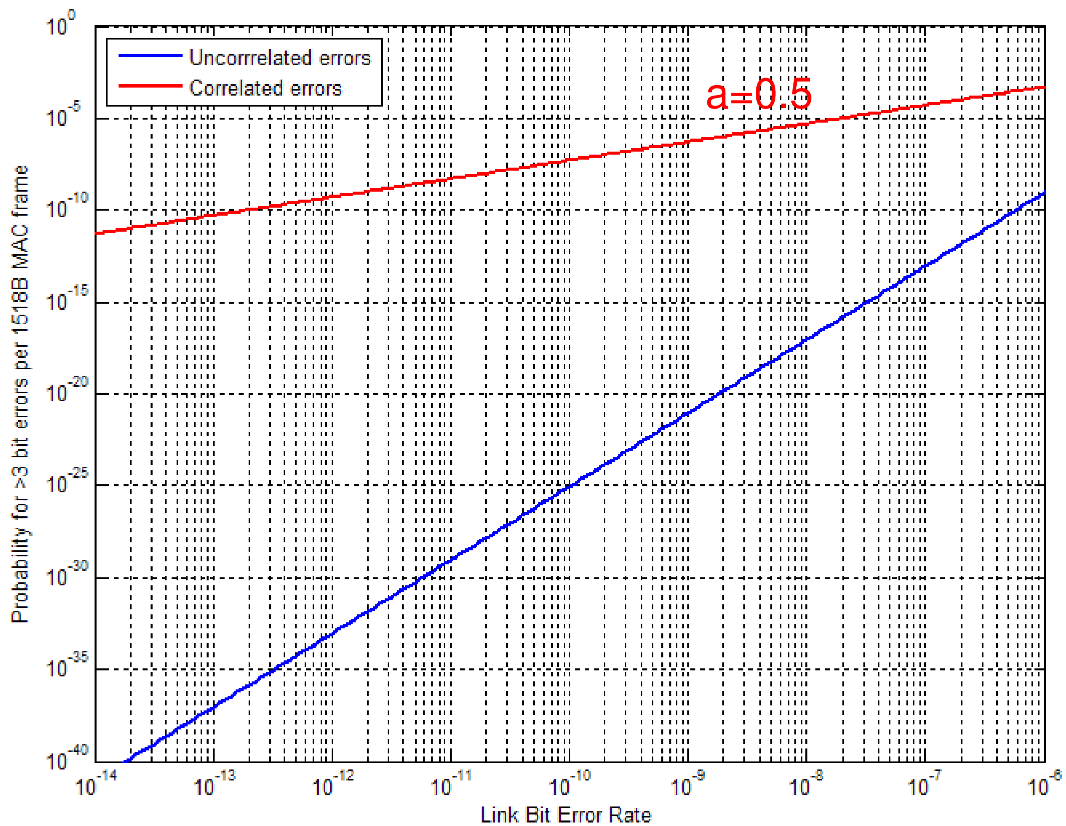
  CRC polynomial: $x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^{8} + x^{7} + x^{5} + x^{4} + x^{2} + x + 1$

- Interaction of scrambler and CRC depends on selected scrambler polynomial: $1 + x^{39} + x^{58}$

- Scrambler polynomial and CRC polynomial do not have a common factor. This is a necessary but not a sufficient condition for not weakening CRC's error detection capability by the descrambler. As the scrambler runs continuously on all payload bits (clause 49.2.6), the effects of error spilling should be considered in CRC performance analysis. An exhaustive search for all error patterns with weight < 4 by explicitly checking all spill-in and spill-out errors showed that CRC's error detection capability is not weakened by 3x error multiplication in the descrambler: *walker_1_0100.pdf*

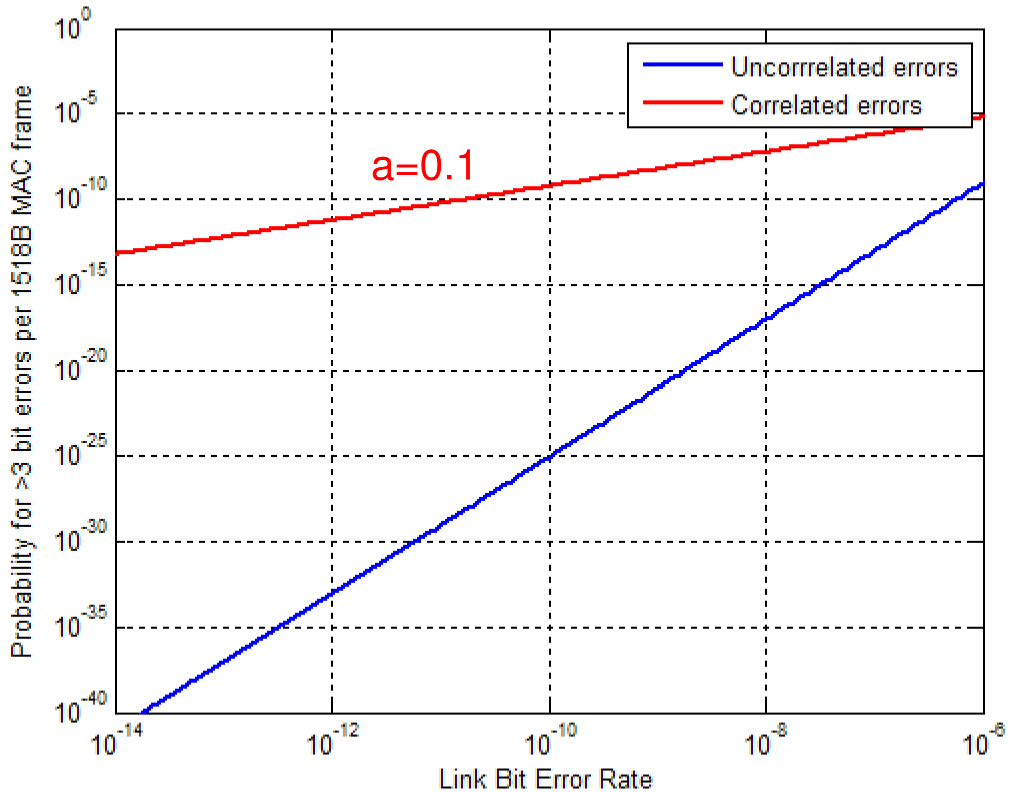
- Compute the probability of having 4 or more link bit errors in a frame to analyze CRC performance in the presence of descrambling. Probability of frames with 4 or more link bit errors that are undetected by CRC is 2^{-32}.

Probability for > 3 Bit Errors – High Correlation Case



- Correlated errors modeled by Gilbert burst error model with prob. of staying in burst state: a = 0.5
- Significant increase in probability for > 3 bit errors in 1518-byte basic ethernet frame in the case of correlated errors

Probability for > 3 Bit Errors – Low Correlation Case



- Correlated errors modeled by Gilbert burst error model with prob. of staying in burst state: a = 0.1
- Significant increase in probability for > 3 bit errors in 1518-byte basic ethernet frame in the case of correlated errors

64b/66b Coding in 100GBASE-R

- 64b/66b coding used in 100GBASE-R (IEEE 802.3ba-2010, Clause 82)
  - 1 type of data block (DB) with 2-bit header **01**
  - 11 types of control blocks (CB) with 2-bit header **10** where the first byte of the payload is rate-4/8 encoded (Hamming distance=4) **8-bit block type field** indicating the type of control block format

	Input Data	Sync	Block Payload							
	Bit Position:	0 1	2							65
	Data Block Format:									
DB	$D_0 D_1 D_2 D_3 D_4 D_5 D_6 D_7$	01	D_0	D_1	D_2	D_3	D_4	D_5	D_6	D_7
	Control Block Formats:		Block Type Field							
CB 1	$C_0 C_1 C_2 C_3 C_4 C_5 C_6 C_7$	10	0x1E	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7
2	$S_0 D_1 D_2 D_3 D_4 D_5 D_6 D_7$	10	0x78	D_1	D_2	D_3	D_4	D_5	D_6	D_7
3	$O_0 D_1 D_2 D_3 Z_4 Z_5 Z_6 Z_7$	10	0x4B	D_1	D_2	D_3	O_0	0x000_0000		
4	$T_0 C_1 C_2 C_3 C_4 C_5 C_6 C_7$	10	0x87		C_1	C_2	C_3	C_4	C_5	C_6	C_7
5	$D_0 T_1 C_2 C_3 C_4 C_5 C_6 C_7$	10	0x99	D_0		C_2	C_3	C_4	C_5	C_6	C_7
6	$D_0 D_1 T_2 C_3 C_4 C_5 C_6 C_7$	10	0xAA	D_0	D_1		C_3	C_4	C_5	C_6	C_7
7	$D_0 D_1 D_2 T_3 C_4 C_5 C_6 C_7$	10	0xB4	D_0	D_1	D_2		C_4	C_5	C_6	C_7
8	$D_0 D_1 D_2 D_3 T_4 C_5 C_6 C_7$	10	0xCC	D_0	D_1	D_2	D_3		C_5	C_6	C_7
9	$D_0 D_1 D_2 D_3 D_4 T_5 C_6 C_7$	10	0xD2	D_0	D_1	D_2	D_3	D_4		C_6	C_7
10	$D_0 D_1 D_2 D_3 D_4 D_5 T_6 C_7$	10	0xE1	D_0	D_1	D_2	D_3	D_4	D_5		C_7
11	$D_0 D_1 D_2 D_3 D_4 D_5 D_6 T_7$	10	0xFF	D_0	D_1	D_2	D_3	D_4	D_5	D_6

Figure 82–5—64B/66B block formats

BTF

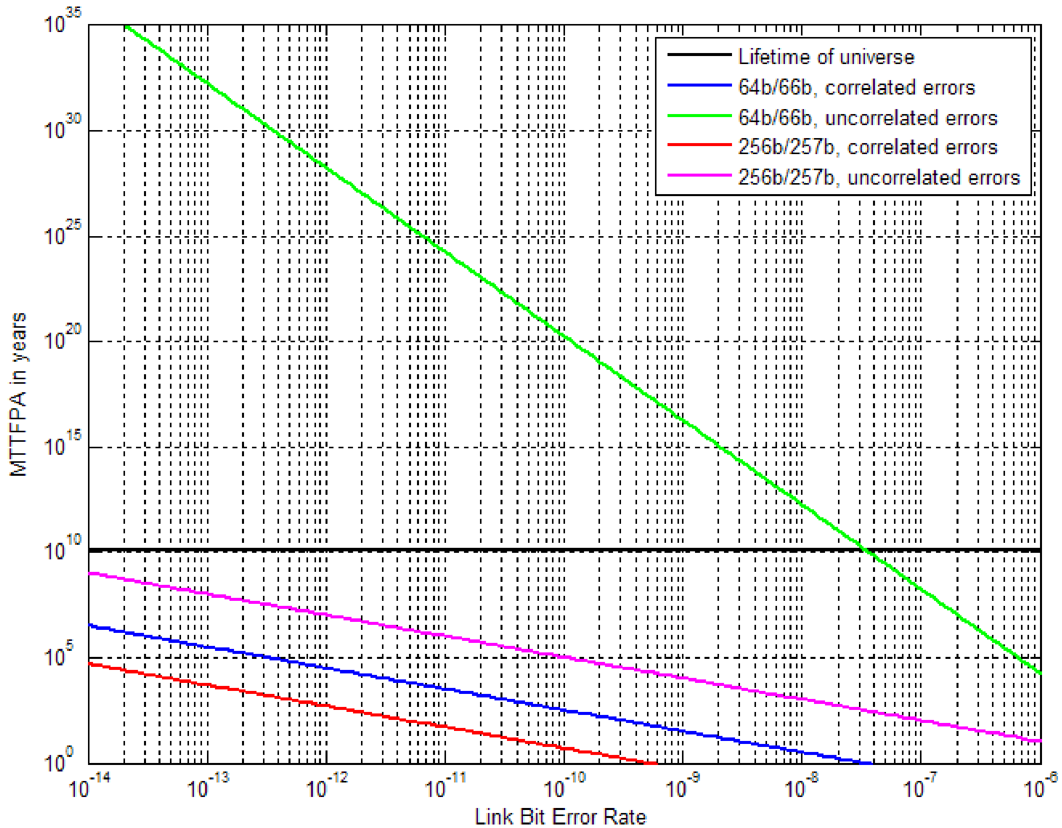
Undetected Errors in Header Bits

- In 100GBASE-R there are 8 possible control blocks indicating frame termination and 1 possible control block indicating frame start
- An error in the header bit of a transcoded block may cause a transcoded block containing control blocks to be recognized as data-only transcoded block
- Assume a transcoded block containing a frame termination control block and a frame start control block is mistaken for a data-only transcoded block as a result of a header bit error.
- In this case two MAC frames can merge into a single large MAC frame.
- If the CRC of the large MAC frame cannot detect the error, an undetected MAC frame error occurs.

MTTFPA Estimation – High Correlation Case

- For uncorrelated errors MTTFPA estimation is similar to MTTFPA of 512b/513b transcoding in *Teshima et al. "Bit-Error-Tolerant (512*N)b/(513*N+1)b Code for 40Gb/s and 100Gb/s Ethernet Transport", INFOCOM 2008*.
- Analysis accounts for correlated errors in MAC frame assuming a=0.5
- Basic ethernet frames of length 1518 bytes
- For correlated errors with a=0.5 MTTFPA is very low both for 64b/66b and 256b/257b

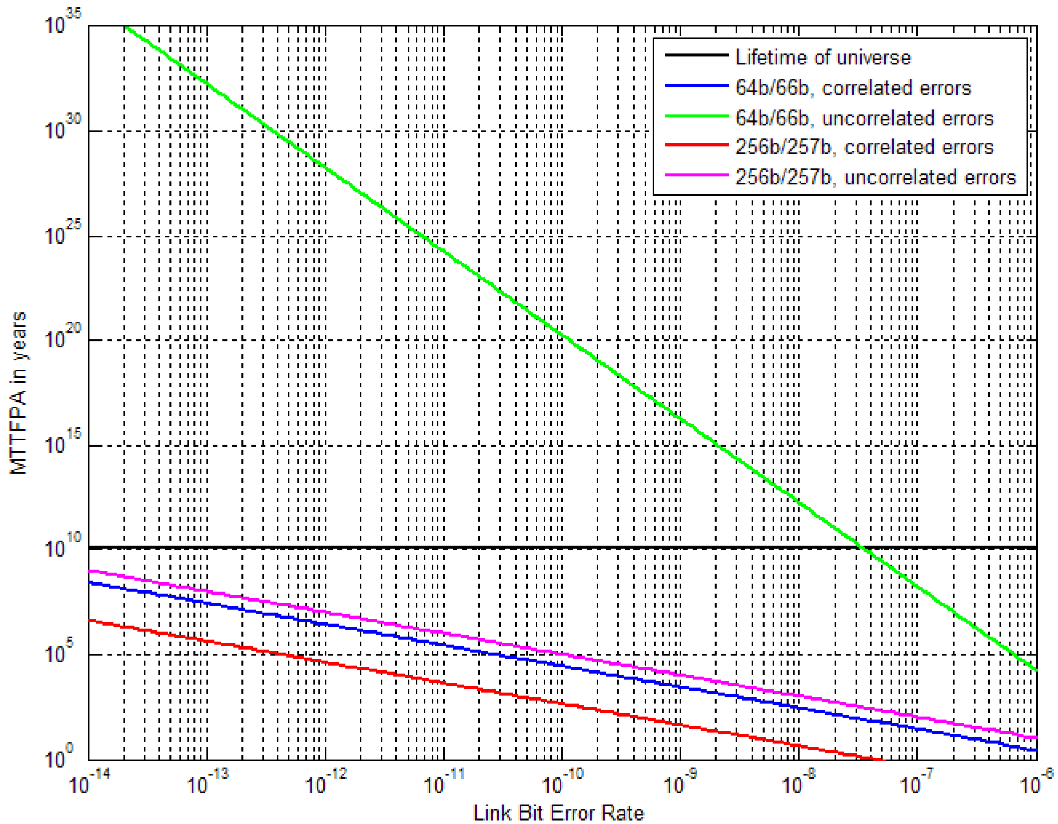
For correlated errors
prob. that burst error continues: a=0.5



MTTFPA Estimation – Low Correlation Case

- Analysis accounts for correlated errors in MAC frame assuming a=0.1
- Basic ethernet frames of length 1518 bytes
- For correlated errors with a=0.1 modest improvement in MTTFPA for 64b/66b and 256b/257b

For correlated errors
prob. that burst error continues: a=0.1



Turning off FEC decoder is not recommended for 256b/257b transcoded and FEC encoded data

Improving MTTFPA for Uncorrelated Errors

- Transcoding (TC) schemes that improve compression: *malpass_03_0907.pdf*
- TC schemes that improve MTTFPA in the case of uncorrelated errors: 1) *Teshima et al. "Bit-Error-Tolerant (512*N)b/(513*N+1)b Code for 40Gb/s and 100Gb/s Ethernet Transport", INFOCOM 2008*, 2) *trowbridge_01_0308.pdf*
- TC schemes with larger transcoded blocks such as 512b/514b, 1024b/1026b and 1024b/1027b achieve good compression and improve MTTFPA in the presence of uncorrelated errors at the expense of an increase in latency and complexity.
- The headers of these schemes are protected by a parity code with Hamming distance 2. Furthermore, all block type fields of control blocks indicating termination of frames have Hamming distance 2 and all block type fields of control blocks indicating start of frames have Hamming distance 2. Fixed-length or variable-length coding are used to achieve Hamming distance 2 in start and termination-type control block type fields.

Improving MTTFPA for Uncorrelated Errors (cont.)

2-bit Hamming distance for control block types with additional compression (flag bit for Remedy 3)

- Flag = 01 – no 66B control blocks in 1026-bit block
- Flag = 10 – One or more 66B control blocks in 1026-bit block
- Flag bits plus sixteen transcoded 66B codewords combined (as below) to form each 1026-bit block
- All 66B control blocks are sorted to front of 1026-bit block with f indicating if there are more control blocks, 4-bit POS indicating original position (row) of this control block
- Minimum 2-bit Hamming distance between coding for block types with terminate /T/ in different positions

Input Data	Block Payload										
Bit Position: / Data Block Format:	0										63
D0D1D2D3/D4D5D6D7	D0			D1	D2	D3	D4	D5	D6	D7	
	Block Type Field										
C0C1C2C3/C4C5C6C7	f	pos	001	C0	C1	C2	C3	C4	C5	C6	C7
S0D1D2D3/D4D5D6D7	f	pos	010	D1	D2	D3	D4	D5	D6	D7	
O0D1D2D3/C4C5C6C7	f	pos	100	D1	D2	D3	O0	C4	C5	C6	C7
T0C1C2C3/C4C5C6C7	f	pos	1100101	C1	C2	C3	C4	C5	C6	C7	
D0T1C2C3/C4C5C6C7	f	pos	1100110	D0	C2	C3	C4	C5	C6	C7	
D0D1T2C3/C4C5C6C7	f	pos	1101001	D0	D1	C3	C4	C5	C6	C7	
D0D1D2T3/C4C5C6C7	f	pos	1101010	D0	D1	D2	C4	C5	C6	C7	
D0D1D2D3/T4C5C6C7	f	pos	10101	D0	D1	D2	D3	C5	C6	C7	
D0D1D2D3/D4T5C6C7	f	pos	10110	D0	D1	D2	D3	D4	C6	C7	
D0D1D2D3/D4D5T6C7	f	pos	000	D0	D1	D2	D3	D4	D5	C7	
D0D1D2D3/D4D5D6T7	f	pos	011	D0	D1	D2	D3	D4	D5	D6	

Minimum Hamming distance 2

NTT Alcatel·Lucent

40GbE MTTFPA when using transcoding - 22 - For IEEE802.3ba, Mar. 2008

trowbridge_01_0308.pdf

Improving MTTFPA for Correlated Errors

- Only 11 type of control blocks in 100GBASE-R as opposed to 15 in 10GBASE_R. Only one type of control block indicating start of frame in 100GBASE-R. Therefore, no need to remap control block type fields indicating start of frame in order to achieve Hamming distance 2.
- TC schemes can be designed such that
  - headers are protected by a parity code with Hamming distance 2
  - all block type fields of control blocks indicating termination of frames have Hamming distance 2
- In the presence of correlated errors we need to interleave header bits and control block type fields along and/or across physical lanes in order to benefit from Hamming distance 2 and improve MTTFPA
- We can design 640b/642b TC where at the beginning of each alignment period instead of 2 transcoded blocks 20 alignment markers and 5 dummy bits are used. RS(528,514)-encoded FEC blocks for NRZ contain in this case 4 dummy bits to ensure 0% overclocking
- Headers and control block type fields indicating termination of frame in 640b/642b TC are protected by fixed-length or variable-length codes with Hamming distance 2. Drawbacks are: 1) about 10 ns increase in latency, 2) significant increase in complexity due to, e.g., variable-length coding and interleaving, 3) improvements in MTTFPA in the case of correlated errors are not sufficient

Summary

- Assuming correlated errors at the DFE output MTTFPA was estimated for uncorrected FEC blocks that have been 256b/257b transcoded
- MTTFPA improvements with alternative transcoding schemes come at the expense of a significant increase in latency and complexity
- Improvements in MTTFPA of uncorrected transcoded FEC blocks are insufficient for NRZ signaling with 0% overclocking
- Transcoding and FEC encoding data without turning on FEC decoder is not recommended
- Significant degradation in MTTFPA for uncoded 64b/66b in the case of correlated errors

Thank You