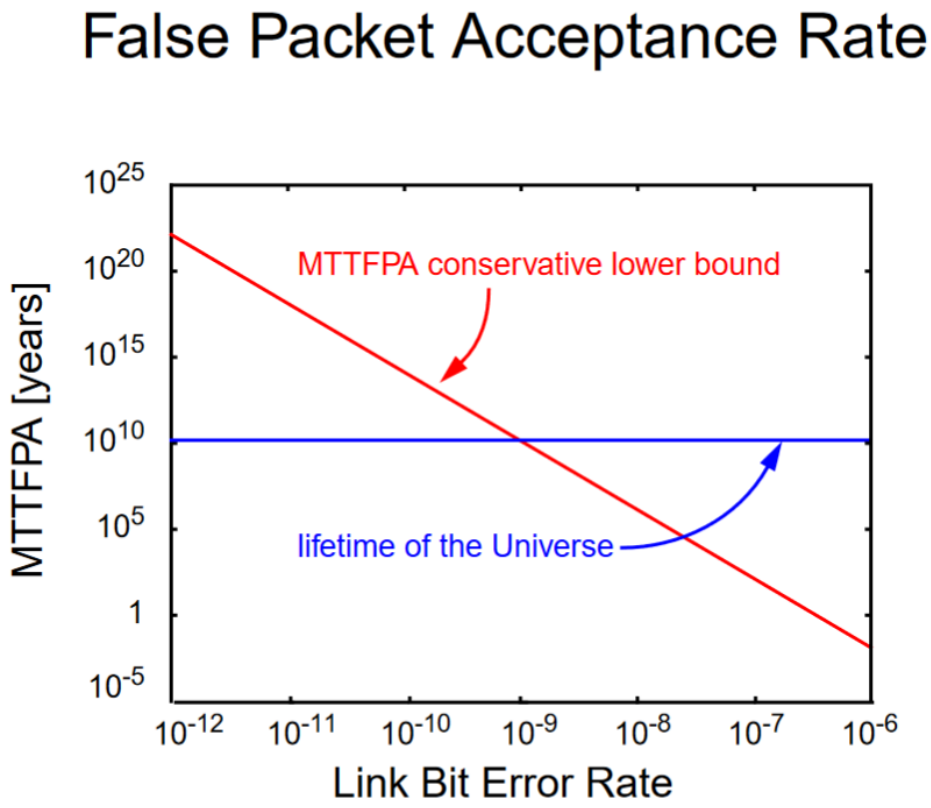
P802.3ct mean time to false packet acceptance (MTTFPA)

Pete Anslow, Ciena

IEEE P802.3ct Task Force, Salt Lake City, May 2019

Introduction

The mean time to false packet acceptance (MTTFPA) has been a key design parameter for many Ethernet projects since P802.3ae (10 Gigabit Ethernet). See walker_1_0300 which predicted a MTTFPA value greater than the age of the universe at the specified link BER:



While this may seem like overkill, users want a low rate of false packet acceptance in a network containing a large number of Ethernet links and also it is desirable that the MTTFPA remains reasonable when the link BER is worse than the specification level.

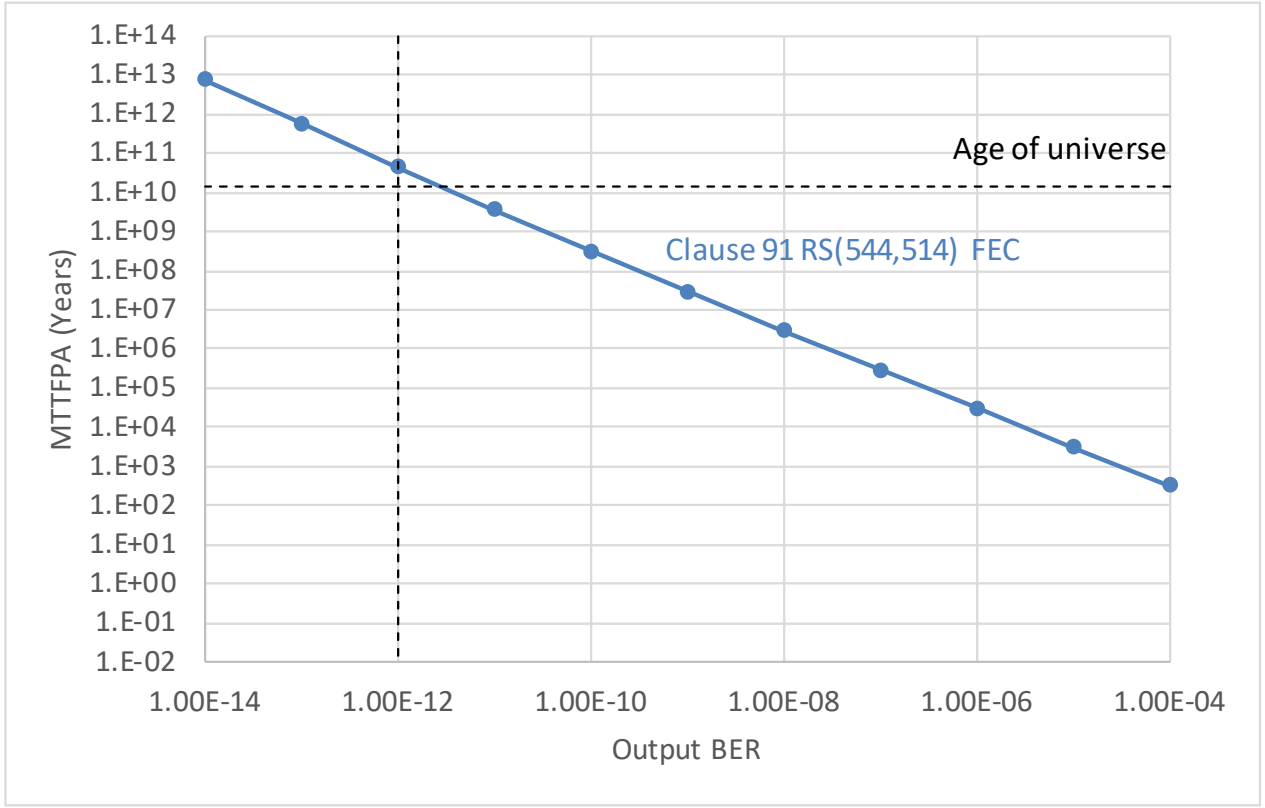
MTTFPA for 100GBASE-ZR

CRC32 and 100GbE 64B/66B PCS

- Ethernet's CRC32 has the following error detection capability
  - All 1, 2 or 3 bit errors are detected
  - All bursts up to 32 bits
  - All double bursts up to 8 bits
  - The above is true for at least up to 9k octet frames
- For the 100GBASE-R scrambler, single bit errors become 3 bit errors - this was shown to not degrade the error detection capability of the IEEE CRC32 for 10GBASE-R
  - No CRC degradation occurs if the CRC and the scrambler polynomial do not share common factors
  - IEEE CRC32 has no common factors with the $X^{\wedge 58}$ scrambler
  - If the original errors can be detected, then the multiplied errors are also detectable
- However, if FEC is used to improve performance, uncorrectable codewords can easily generate 4 or more errors in a single Ethernet frame

Clause 91 RS(544,514) MTTFPA performance

Clause 91 requires that the FEC decoder “shall also be capable of indicating when an errored codeword was not corrected. The probability that the decoder fails to indicate a codeword with *t*+1 errors as uncorrected is not expected to exceed 10^{-6}. This limit is also expected to apply for *t*+2 errors, *t*+3 errors, and so on.”



This means that the Ethernet frames contained in nearly all uncorrected codewords are replaced with Error blocks. The remaining errored frames are mostly discarded by the MAC due to the CRC32 resulting in the MTTFPA curve shown here.

100GBASE-ZR FEC scheme

The adopted baseline for the 100GBASE-ZR FEC and frame format is on slides 9 to16 of trowbridge_3cn_01a_0119. This is:

- A FEC frame similar to the OTU4 frame specified in ITU-T G.709
- “Staircase” FEC as specified in ITU-T G.709.2 Annex A
- A GMP mapper as defined in ITU-T G.709 that inserts the serialized and deskewed stream of 66B blocks into the Staircase FEC frame
- A frame synchronous scrambler (so no error extension)

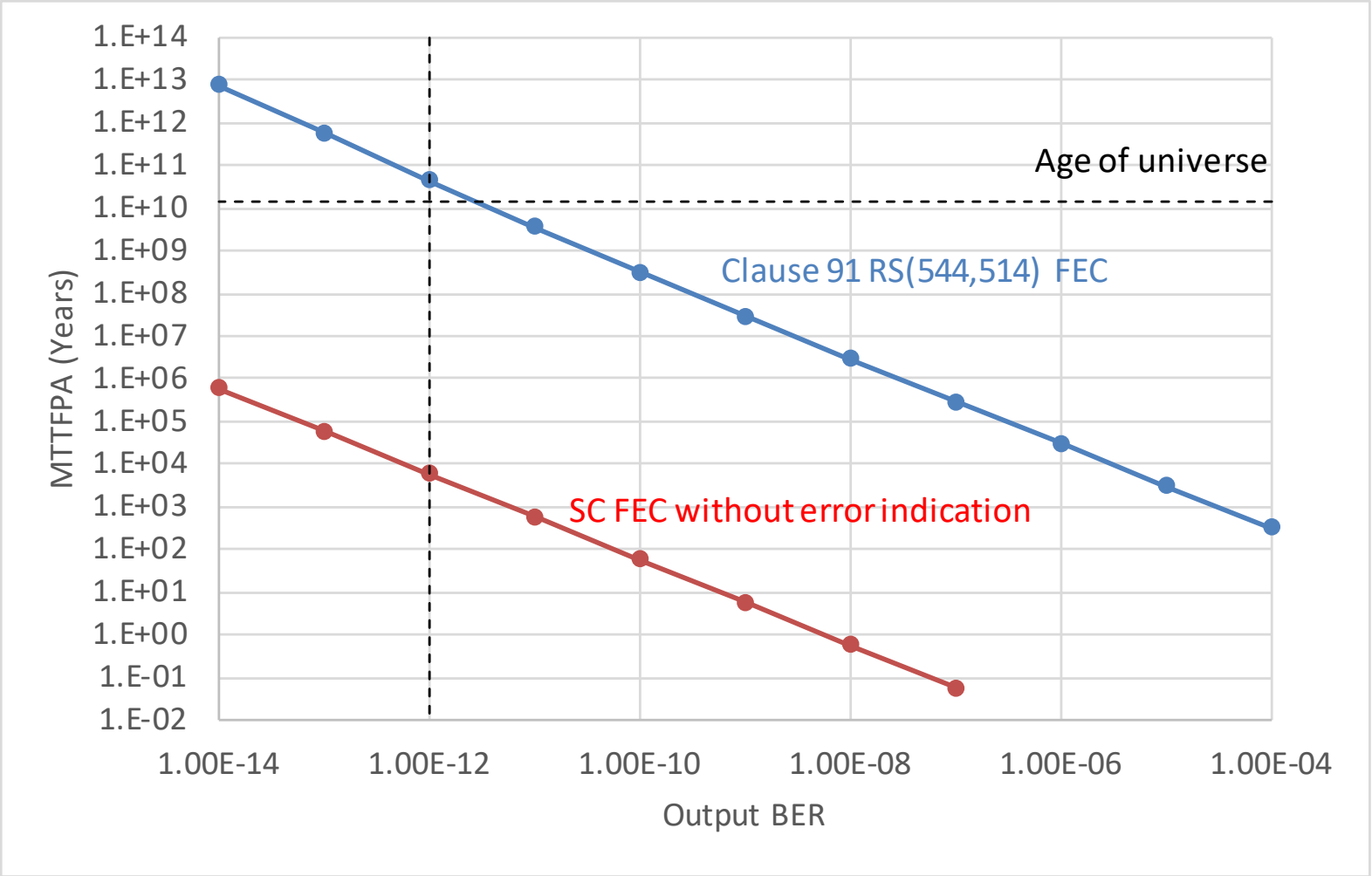
However, there is no requirement so far that the Ethernet frames inside uncorrectable FEC blocks are marked as bad.

Assumptions about Staircase FEC for 100GBASE-ZR

In order to assess the MTTFPA for 100GBASE-ZR, the following assumptions have been made:

- Ethernet frames inside uncorrectable FEC blocks are **not** marked as bad.
- The error distribution within an uncorrectable FEC block at the output is similar to the error distribution within that block at the input.
- The pre-FEC BER for staircase FEC is as per Table A.1 in ITU-T G.709.2
- The input error distribution is random and not bursty.
- All Ethernet frames have the worst case length for MTTFPA
- The MAC fails to drop errored Ethernet frames due to a false CRC32 match with a probability of 2.33E-10.

100GBASE-ZR MTTFPA performance



Without error indication, the MTTFPA is about 7 orders of magnitude worse than for Clause 91 RS(544,514), e.g., 100GBASE-SR2

Conclusion for 100GBASE-ZR

The mean time to false packet acceptance (MTTFPA) performance shown on the previous slide is far below the level expected for Ethernet PHYs.

Consequently, some change has to be made to the FEC definition to improve the MTTFPA.

One solution would be to mark some of the 66B blocks contained in uncorrectable FEC blocks as invalid (in a similar manner as for Clause 91), while avoiding exceeding the threshold for hi_ber.

MTTFPA for 400GBASE-ZR

400GBASE-ZR FEC scheme

The adopted baseline for the 400GBASE-ZR PCS/PMA is in lyubomirsky_3cn_01b_0119. This is:

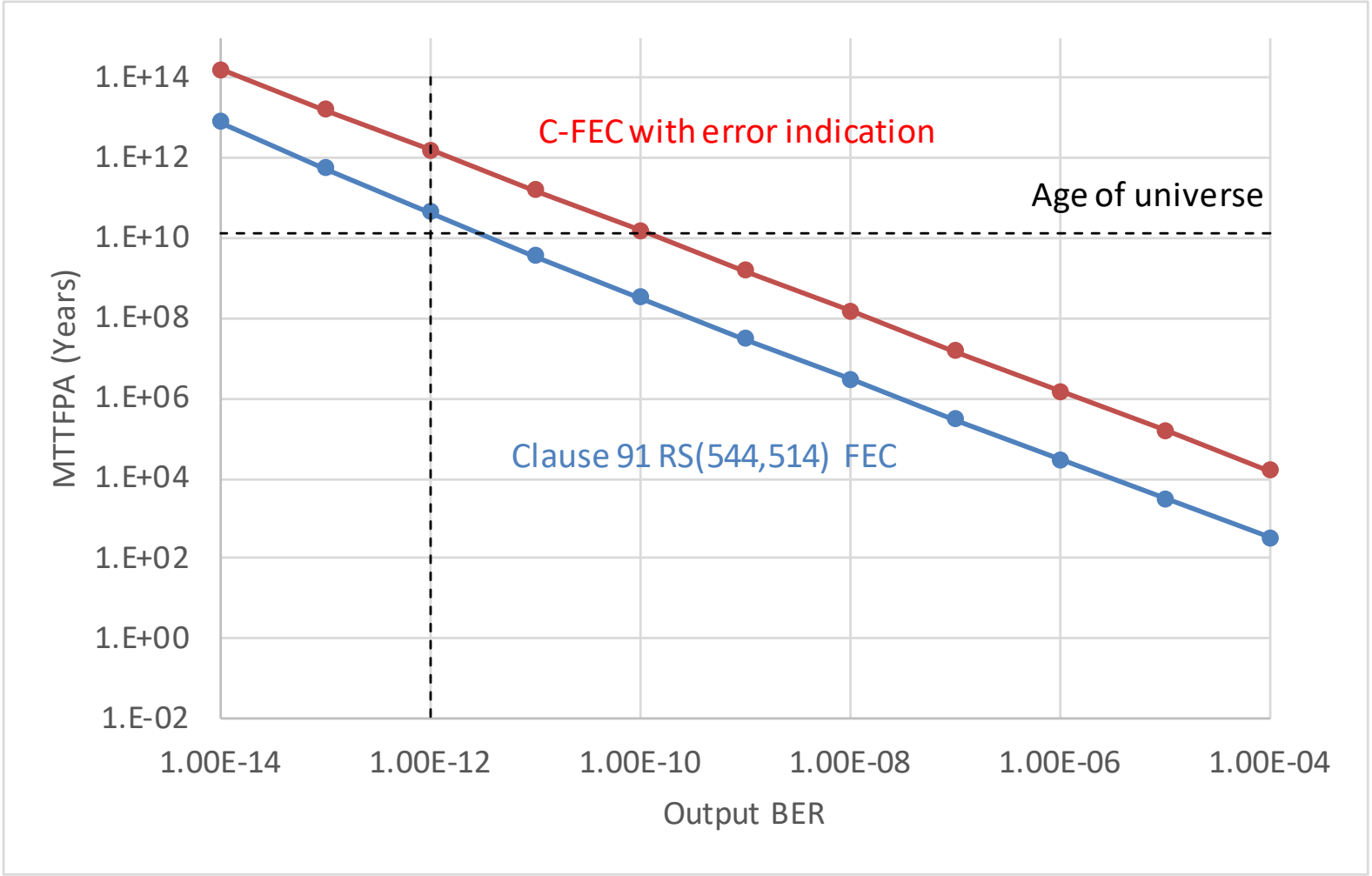
- A frame similar to the 400ZR frame specified in the draft OIF 400ZR IA
- C-FEC – “Staircase” FEC as specified in ITU-T G.709.2 Annex A as an outer code and a (128,119) Hamming soft decision decoding inner code
- A GMP mapper as defined in ITU-T G.709 that inserts a 256b/257b transcoded stream of 66B blocks into the frame
- A frame synchronous scrambler (so no error extension)
- An additional CRC32 added to the frame. If this does not match or the C-FEC decoder detects uncorrected errors, the whole block (952 x 257-bit transcoded blocks) is replaced by transcoded 66B error control blocks.

Assumptions about C-FEC for 400GBASE-ZR

In order to assess the MTTFPA for 400GBASE-ZR, the following assumptions have been made:

- Ethernet frames inside uncorrectable FEC blocks **are** marked as bad.
- The error distribution within an uncorrectable FEC block at the output is similar to the error distribution within that block at the input (probably conservative).
- The pre-FEC BER for C-FEC is as per Table II.2 in ITU-T G.709.3
- The input error distribution is random and not bursty.
- All Ethernet frames have the worst case length for MTTFPA
- The FEC decoder fails to replace uncorrectable FEC blocks due to a false CRC32 match with a probability of 2.33E-10.
- The MAC fails to drop errored Ethernet frames due to a false CRC32 match with a probability of 2.33E-10.

400GBASE-ZR MTTFPA performance



With error indication, the MTTFPA is about 1 order of magnitude better than for Clause 91 RS(544,514), e.g., 100GBASE-SR2

Conclusion for 400GBASE-ZR

The mean time to false packet acceptance (MTTFPA) performance shown on the previous slide is above the level expected for Ethernet PHYs.

Consequently, as long as the probability of a false match to one of the CRC32's remains at 2.33E-10 even when there is a false match to the other one, no further action is required.

Thanks!