BER objective format for 400GbE

Pete Anslow, Ciena

IEEE 400 Gb/s Ethernet Study Group, Logic Ad Hoc, 26 June 2013

Introduction

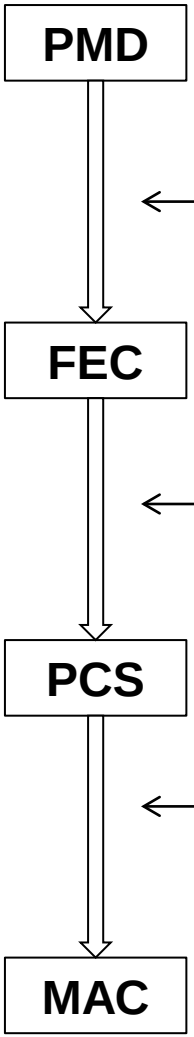
The error performance objective adopted for the P802.3ba, P802.3bj and P802.3bm projects was:

“Support a BER better than or equal to 10^{-12} at the MAC/PLS service interface”

However, when it was decided to employ FEC for most of the new PHYs in P802.3bj and P802.3bm, this objective could no longer be directly applied since we need far fewer unmarked errors than this at the MAC/PLS service interface in order to meet MTTFPA (Mean Time To False Packet Acceptance) expectations.

This contribution is aimed at defining an alternative performance objective format that is directly applicable for PHYs with and without FEC.

Flow through P802.3bj FEC enabled stack



The BER at the FEC input may be much higher than the PHY performance objective. The BER required to meet the objective depends on the error statistics.

Correctable errors have been corrected (unless correction is bypassed). Detected but uncorrected errors are marked as bad using sync header violations.

Some 66B blocks from FEC codewords containing detected but uncorrected errors have been converted to /E/ control codes. The only errors present but not marked are undetected errors which are very rare.

MAC frames missing their start or terminate control codes or containing /E/ control codes or with invalid CRC are discarded.

BER at the MAC/PLS service interface

As shown on the previous slide, at the MAC/PLS service interface (just above the MAC on the diagram on the left) the BER is very low in this FEC enabled architecture. The only errored bits are those that were not detected by the FEC decoder.

We can get an estimate as to how often an error appears at this point in the stack from the MTTFPA target of the age of the universe.

The FEC scheme proposed to be used for 100GBASE-CR4/KR4/SR4 is capable of correcting all error patterns in a FEC codeword containing 7 or less errored symbols. This means that when a FEC codeword contains any undetected errors, there must be at least 8 of them. However, the CRC used by Ethernet frames is only capable of guaranteed detection of up to 3 errored bits located anywhere in a frame. For more errors than this it has a probability of failing to detect errors of 2^{-32}. This means that a frame containing errors can only arrive at the MAC every 13.8E9/2^32 = 3.2 years.

Effect of uncorrectable errors

For the stack shown on slide 3, the dominant effect of uncorrected errors at the FEC output is not that errors appear at the MAC/PLS service interface, it is that frames are discarded.

However, this is also true for 64B/66B coded Ethernet systems without FEC. Here, nearly all errored frames contain 3 or less errors and are guaranteed to be discarded by the MAC because the CRC does not match the data. (Errored frames not guaranteed to be discarded only arrive once every 3 years).

This means that if we set the error performance objective as a minimum Frame Loss Ratio (FLR), then this can be directly applied to both 64B/66B coded and FEC enabled PHYs.

This is in accordance with the resolution of Comment #42 against P802.3bj D2.0 which has defined performance using:

> frame loss ratio (the number of transmitted frames not received as valid by the MAC divided by the total number of transmitted frames) for 64-octet frames with minimum inter-packet gap.

What is the relationship between BER and FLR?

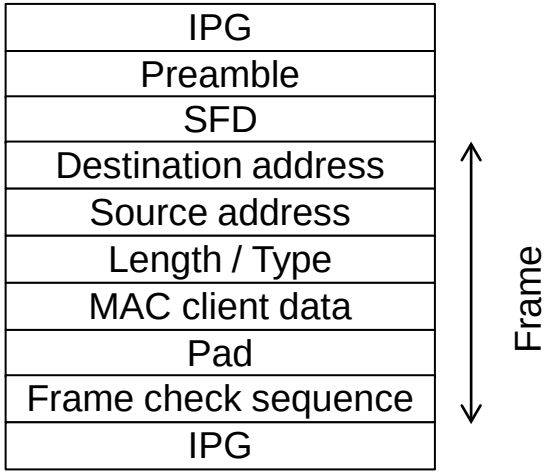
For the P802.3ba project the objective of a BER of better than or equal to 10^{-12} at the MAC/PLS service interface resulted in the BER at the PMD service interface being required to be better than or equal to 10^{-12}

For the P802.3bj and P802.3bm projects the error performance objective was still defined as a BER. For FEC enabled applications this was then translated into an FLR requirement by calculating what FLR would result from that BER at the FEC decoder output in a 64B/66B coded system.

Consequently, this contribution proposes to follow the same principle for the 400GbE project and set the FLR objective by calculating what FLR would result from the desired BER at the FEC decoder output in a 64B/66B coded system.

Size of MAC frames after 64B/66B coding

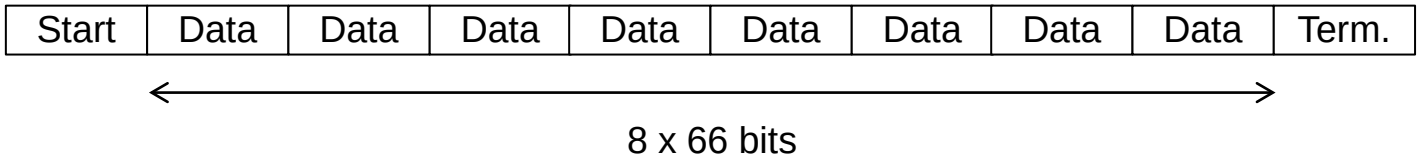
A MAC frame starts with the Destination Address and ends with the frame check sequence. These bits are preceded by the interpacket gap (IPG), 7 octets of preamble and 1 octet of start-of-frame delimiter (SFD).



The first octet of the preamble is mapped to a start control character by the RS and is always aligned to the start of a 64-bit block. Consequently, a 64 octet frame will be encoded as a Start 66-bit block (which contains the Preamble and SFD), followed by eight 66-bit blocks containing the MAC frame, followed by a Terminate 66-bit block containing 7 Idle control characters – 10 66-bit blocks in all with minimum interpacket gap.

Errors causing a frame to be dropped

As described on the previous slide, a 64 octet MAC frame with minimum interpacket gap after 64B/66B coding is a Start block, 8 data blocks and a terminate block.



According to the definition of “R_TYPE” in 82.2.18.2.3, Start is recognised as “a sync header of 10 and a block type field of 0x78” and Terminate is recognised as “a sync header of 10, a block type field of 0x87, 0x99, 0xAA, 0xB4, 0xCC, 0xD2, 0xE1 or 0xFF and all control characters are valid”

Therefore, with 64B/66B coding a frame will be dropped if there is an error in 8 x 66 bits for the data blocks + 10 bits in the Start block + 66 bits for the terminate block = 604 bits. Because of the error multiplication in the descrambler, it will also be dropped if there were errors in 16 of the preceding 58 bits, making a total of 620 bits that must be correct at the descrambler input per frame.

FLR from BER

If we assume that the errors are randomly distributed, then the FLR (as defined on page 5) can be found from:

$$FLR = 1-(1-BER)^{620} \quad (1)$$

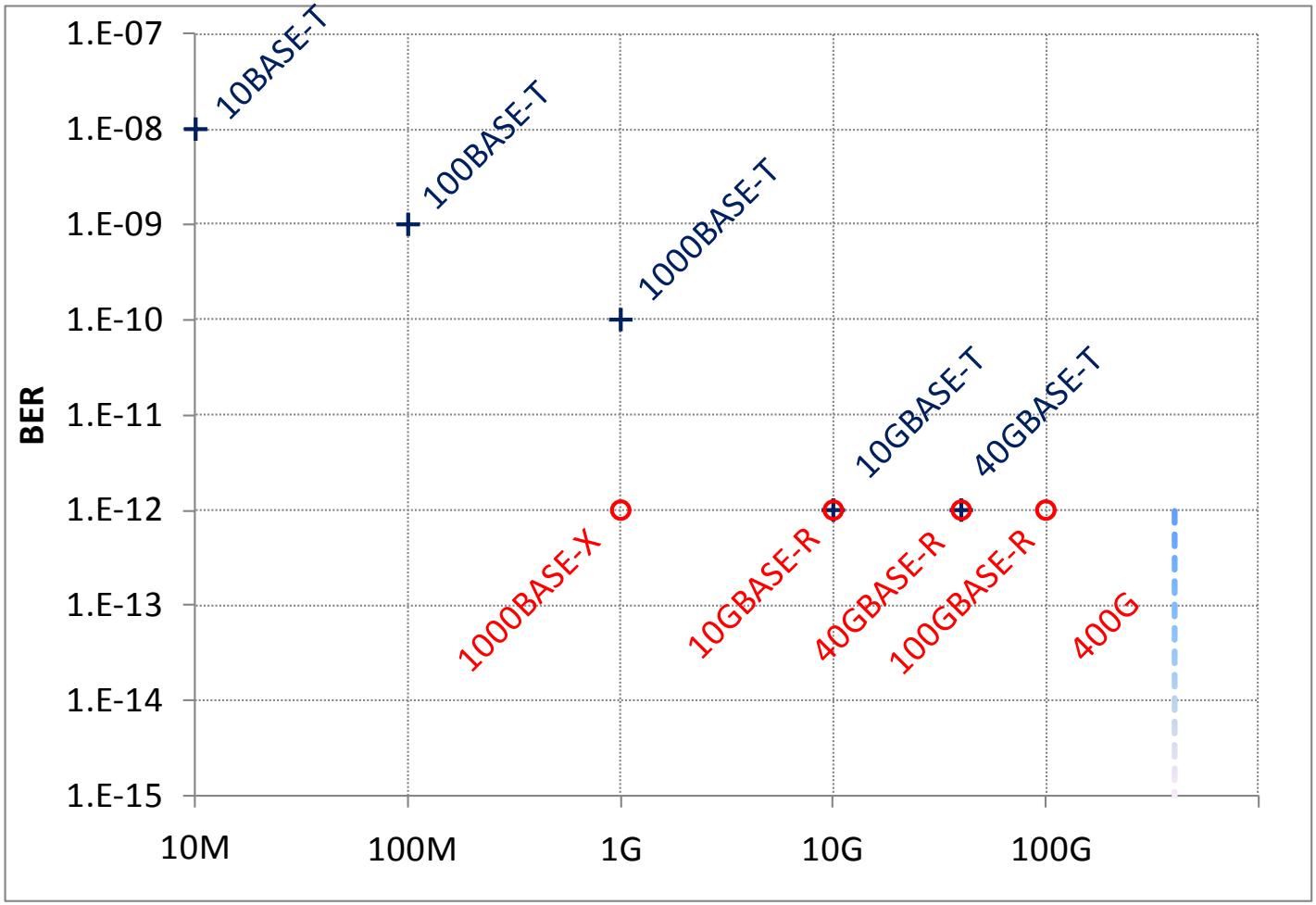
For BER in the range of interest, this can be approximated by:

$$FLR = BER * 620 \quad (2)$$

For BERs that might be candidates for the 400GbE objective, this is:

BER	FER
10^{-12}	6.2×10^{-10}
10^{-13}	6.2×10^{-11}
10^{-14}	6.2×10^{-12}
10^{-15}	6.2×10^{-13}

Ethernet BER vs. bit rate



Conclusion

Since it is fairly likely that at least some 400GbE PHYs will incorporate FEC, it is proposed to set the error performance objective in the form:

“Support a frame loss ratio better than or equal to 6.2×10^{-x}”

Thanks!