Low Latency Reed Solomon Forward Error Correction

Web: http://25gethernet.org/

Revision History

Revision	Who	Date	Change Description
0.1	RJS/BRCM	10/19/2017	Initial Draft Release
0.2	RJS /BRCM	12/19/2017	Added detail on Autonegotiation, clarification on AM spacing
1.0	RJS / BRCM	11/9/2018	Codeword examples added, final revision

Table of Contents

List of Figures

1 Overview

This specification provides functional details for instantiation of forward error correction based on a shortened Reed Solomon code, RS(272).

2 Standards Reference

References are made throughout this document to IEEE 802.3-2015 Ethernet Access Method and Physical Layer [base standards]. It should be noted that at the time of publication of this

specification, 802.3bs and 802.3cd were in draft status although at a state of high maturity, so the clauses in these drafts are also referenced.

- Clause 91 - Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 100GBASE-R PHYs
- Clause 119 - Physical Coding Sublayer (PCS) for 64B/66B, type 200GBASE-R and 400GBASE-R
- Clause 134 - Reed-Solomon Forward Error Correction (RS-FEC) sublayer for 50GBASE-R PHYs

Note: IEEE 802 (e.g. IEEE 802.3-2015, etc) standards documents are available free through IEEE's Get program including IEEE 802 from http://standards.ieee.org/about/get/, six month after publication of each.

3 Low Latency Forward Error Correction (LL-FEC)

In an effort to improve end to end latency, a low-latency shortened codeword variant of the 802.3 FEC (LL-FEC) may OPTIONALLY be used in place of the IEEE 802.3bs and 802.3cd standard RS(544). Following the conventional nomenclature, the LL-FEC code used is RS(n = 272, k = 257+1[1], t = 7, m = 10) where n is the codeword length, k is the message length, t is the maximum number of symbols able to be corrected per codeword, and m is the symbol size (in bits).

For 50 and 100GE ports the LL-FEC substitutes IEEE 802.3cd Clause 134. For 200GE ports the LL-FEC substitutes the FEC specifications contained within IEEE 802.3bs Clause 119.

As the codeword length is half the codeword length of RS(544), it expected that the overall latency associated with error correction will scale accordingly.

The LL-FEC codeword contains 272 x 10 bit symbols (rather than the 544 x 10 bit symbols specified in 802.3bs and 802.3cd). The symbol distribution process from the output of the encoder to the FEC lanes remains the same as 802.3bs or 802.3cd, however the codeword length is 272 symbols rather than the 802.3 specified 544.

Note: as the code rate of 272 codeword output symbols per 257 input (message) symbols is identical to RS(544), the code rate remains the same in both RS(272) and RS(544) cases.

[1] Here the 257+1 nomenclature is used to indicate that the useful message size is in fact 257 symbols, with a single pad symbol used to enable rate and logic commonality with the 802.3 specified RS(544).

3.1 LL-FEC Details

3.1.1 Generator Polynomial

The Generating polynomial used for all LL-FEC operational modes is identical to the one used IEEE 802.3, Eq. 91-1 (or identically 802.3bs, Eq. 119-1), using the using the RS(528,514) coefficients in IEEE 802.3 table 91-1.

3.1.2 Message Polynomial

The message polynomial used for all LL-FEC operational modes is identical to the one used in IEEE 802.3 Eq. 91-2 (or identically IEEE 802.3bs, Eq. 119-2).

3.1.3 Parity Polynomial

The parity polynomial is calculated in identical manner to as specified in IEEE 802.3 Eq. 91-3 (or identically IEEE 802.3bs, Eq. 119-3).

3.1.4 FEC Input Symbols

The input symbols from the transcoder are derived from transcoding 10 sets of 4 x 66 bit PCS blocks into 10 sets of 257 bit transcoding blocks. The resultant output is 2570 bits with a 10 bit pad symbol (see 3.2) which constitutes the FEC input message of 2580 bits total.

3.2 10-bit Pad Symbol

In order to ensure physical layer and logic commonality with the existing 802.3bs and 802.3cd standards, the input to output symbol ratio for LL-FEC is the same as RS(544) (i.e. 544/514 = 272 / 257). However, as RS codes require an even number of parity symbols, a 10 bit pad symbol is added to the end of the message which effectively makes the transcoded data passed to the FEC encoder 258 symbols in length.

The pad symbol may be derived from a free-running PRBS9 pattern, or programmed to a fixed value. The contents of the pad symbol are RECOMMENDED to ensure clock content and baseline wander requirements are met. The pad forms the last symbol of the 258 symbol message to be transmitted and is included in the parity calculation; however it is discarded subsequent to error correction after receipt.

3.3 Symbol Distribution

After FEC encoding, the codewords are distributed in an identical manner as specified in 802.3cd (for 50 and 100GE) or 802.3bs (for 200GE), with the important difference that the LL-FEC codeword length is 272 x 10 bit symbols rather than the standard 544. The transmit bit ordering and distribution are shown below in Figure 1 - Figure 3. A pad symbol is added to each

message prior to the RS-encoder. It also may be instructive to the reader to compare to the analogous IEEE 802.3 figures 91-6, 119-10, 119-11, and 134-4.

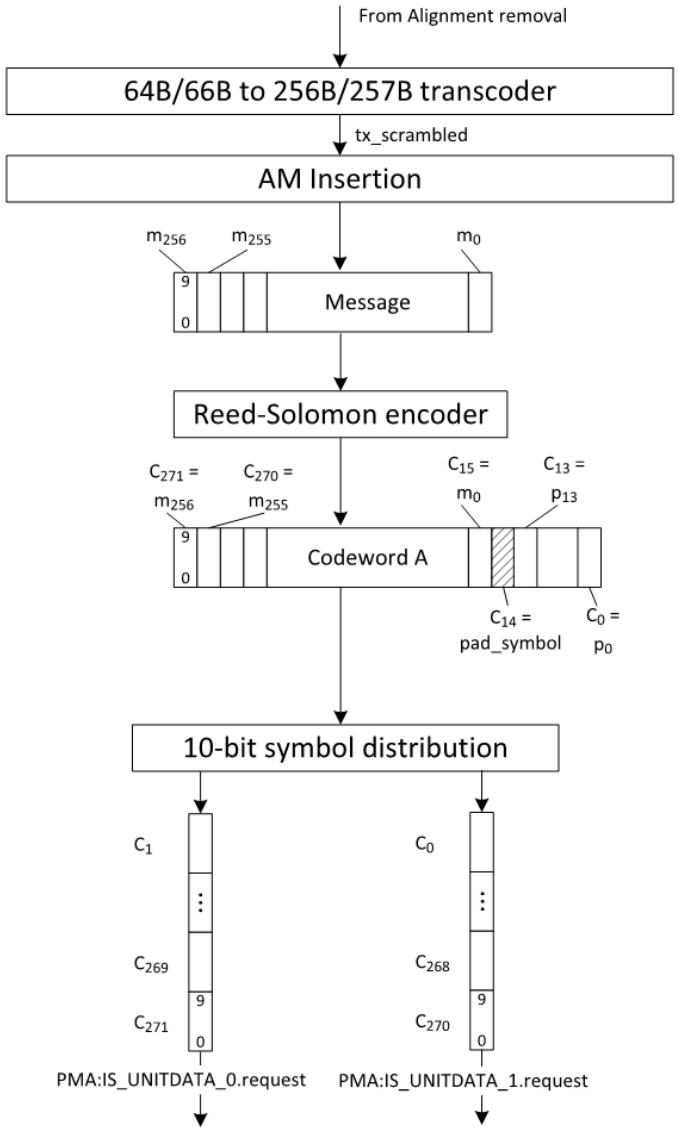

Figure 1 - 50GE Transmit Bit Ordering

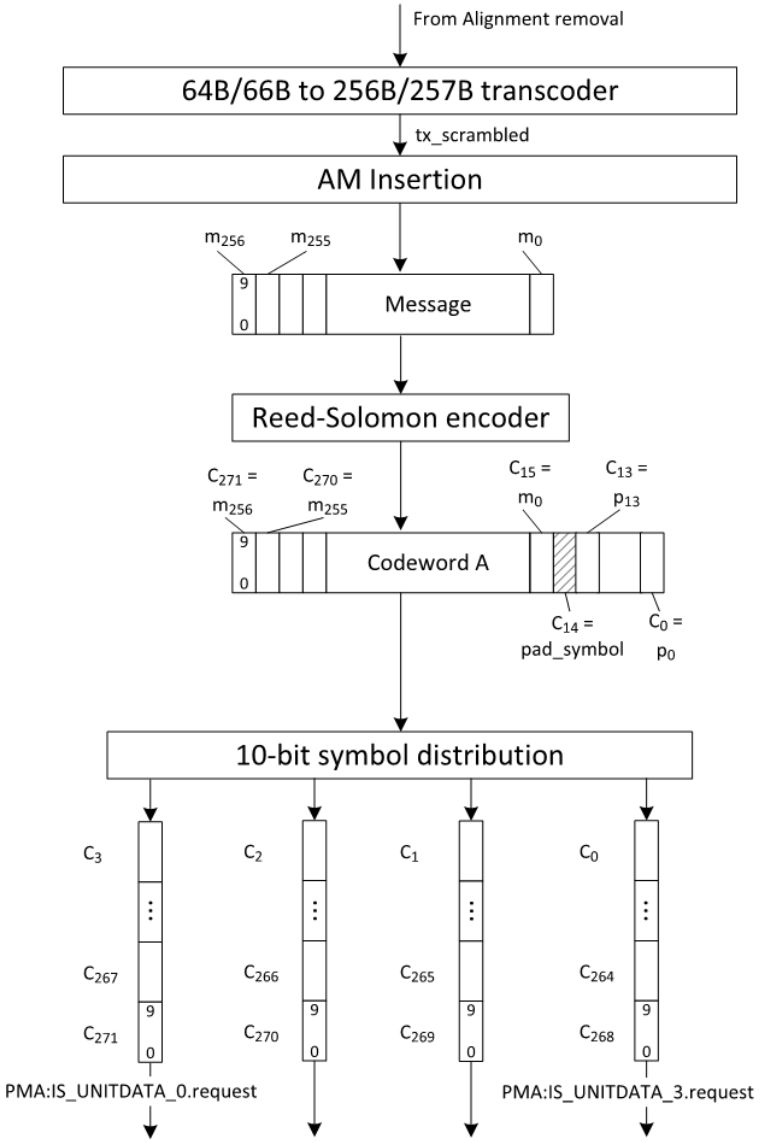

Figure 2 - 100GE Transmit Bit Ordering

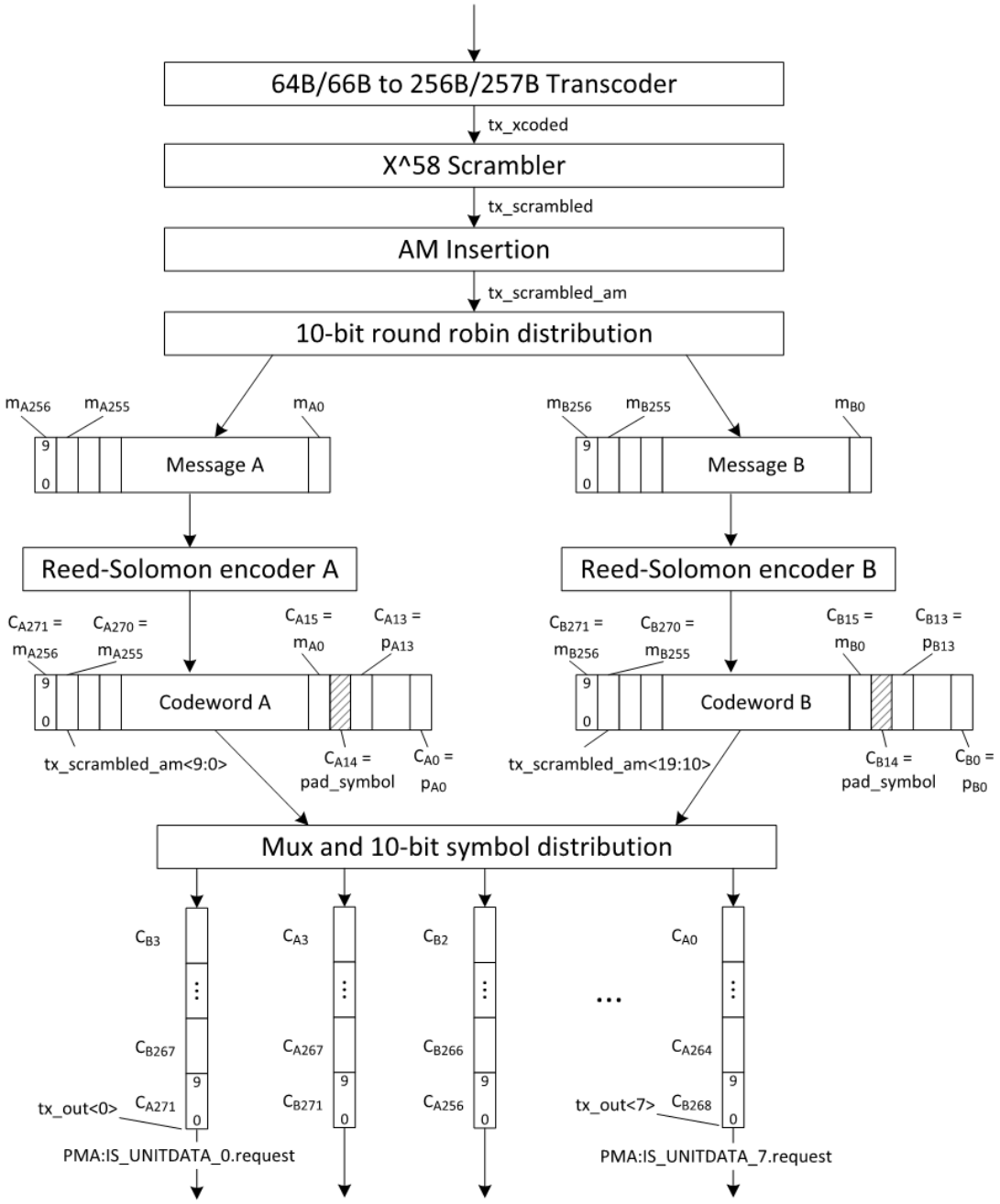

Figure 3 - 200GE Transmit Bit Ordering

3.4 Alignment Marker Insertion Period

For 50 and 100GE LL-FEC operation, as the alignment marker insertion occurs above the FEC sublayer, there is no change to the alignment marker insertion process.

For 200GE LL-FEC operation the alignment marker insertion period in terms of 257-bit transcoding blocks is unchanged from 802.3bs. Alignment markers must be inserted every 81920 x 257-bit transcoding blocks. However as the LL-FEC RS(272, 257+1) codeword is half the length of the RS(544,514) codeword used in 802.3bs, the alignment marker insertion period in terms of codewords is scaled accordingly so that alignment markers are inserted every 8192 FEC codewords (rather than 4096 as specified in 802.3bs).

3.5 Autonegotiation

The LL-FEC uses the existing 25GE Consortium Next Page autonegotiation mechanism, adding provision for ability and request bits as follows (refer also to the 25GE Consortium Specification, Schedule 3, Section 3.2.5 for further details).

3.5.1 Consortium Next Page Format

The 25GE Consortium OUI tagged unformatted Next Page is shown below in

Figure 4. (Note that the formatted message page 5 remains the same as per Schedule 3, 3.2.5.)

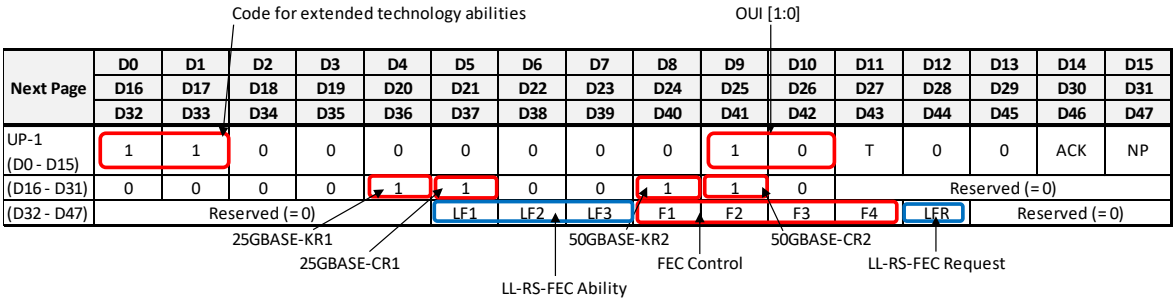

Figure 4 - Unformatted OUI Next Page, Showing Additional Fields for LL-FEC

The new fields added for control of the LL-FEC are shown outlined in blue:

- D37 (LF1): LL FEC ability for 50GBASE-CR1 / KR1
- D38: (LF2) LL FEC ability for 100GBASE-CR2 / KR2
- D39: (LF3) LL FEC ability for 200GBASE-CR4 / KR4

- D44: (LFR) LL FEC request bit

3.5.2 Autonegotiation Resolution

Whether a link partner should use LL-FEC is determined by processing a combination of the Consortium next page bits, as well as the IEEE base page according to the following pseudocode:

Where HCD is the IEEE HCD determined by base page priority, LD = Local Device and RD = Remote Device

```
Where HCD is the IEEE HCD determined by base page priority, LD = Local
Device and RD = Remote Device

IF ((HCD == IEEE 200GBASE-KR4/CR4) & (LD.LF3 & RD.LF3) & (LD.LFR |
RD.LFR)) THEN operate as 200GBASE-KR4/CR4, substituting RS(272) FEC
for RS(544) FEC

IF ((HCD == IEEE 100GBASE-KR2/CR2) & (LD.LF2 & RD.LF2) & (LD.LFR |
RD.LFR)) THEN operate as 100GBASE-KR2/CR2, substituting RS(272) FEC
for RS(544) FEC

IF ((HCD == IEEE 50GBASE-KR1/CR1) & (LD.LF1 & RD.LF1) & (LD.LFR |
RD.LFR)) THEN operate as 50GBASE-KR1/CR1, substituting RS(272) FEC for
RS(544) FEC
```

Note: per IEEE, autonegotiation only applies to operation over copper cable and backplane channels.

3.6 Annex A [Informative]: Error Correcting Performance of RS(272)

CAUTION: The error correcting performance of RS(272) LL-FEC is presented below as compared to RS(544). It should be noted that as the correction capability both in terms of raw gain in presence of random errors, as well as burst error protection is less than the 802.3 standard RS(544). Therefore implementers are strongly recommended to perform their own signal integrity and link modeling validation before deployment of systems based on LL-FEC.

3.6.1 Performance of LL-FEC with 50, 100GE PHYs

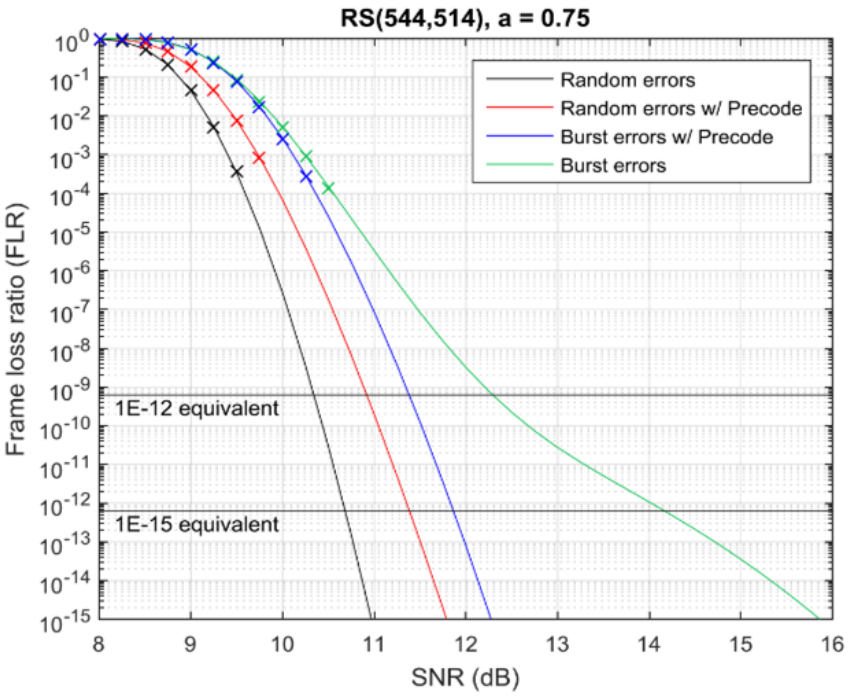

Figure 5 - RS(544) Single Part Link Performance (50, 100GE)

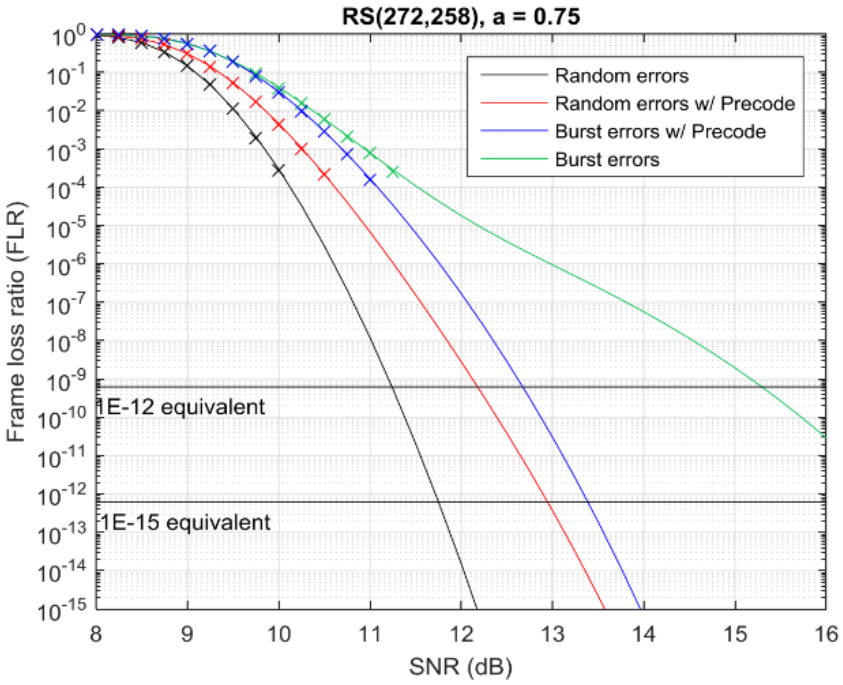

Figure 6 - RS(272) Single Part Link Performance (50, 100GE)

Table 1- 50 & 100GE FEC Performance for RS(544) and RS(272)

FLR = 6.2E-10	RS(544,514)		RS(272,258)	
	BER	SNR (dB)	BER	SNR (dB)
Random errors	3.7677e-04	10.3438	9.9248e-05	11.2402
Random errors w/ precode	1.6504e-04	10.9219	1.8067e-05	12.1797
Burst errors, a=0.75	5.7878e-05	12.2968	8.8834e-09	15.2969
Burst errors w/ precode, a=0.75	3.1334e-04	11.3828	2.5255e-05	12.6758

Note: required BER is reported at slicer output, and includes x 4 when bursts are present

3.6.2 Performance of LL-FEC with 200GE PHYs

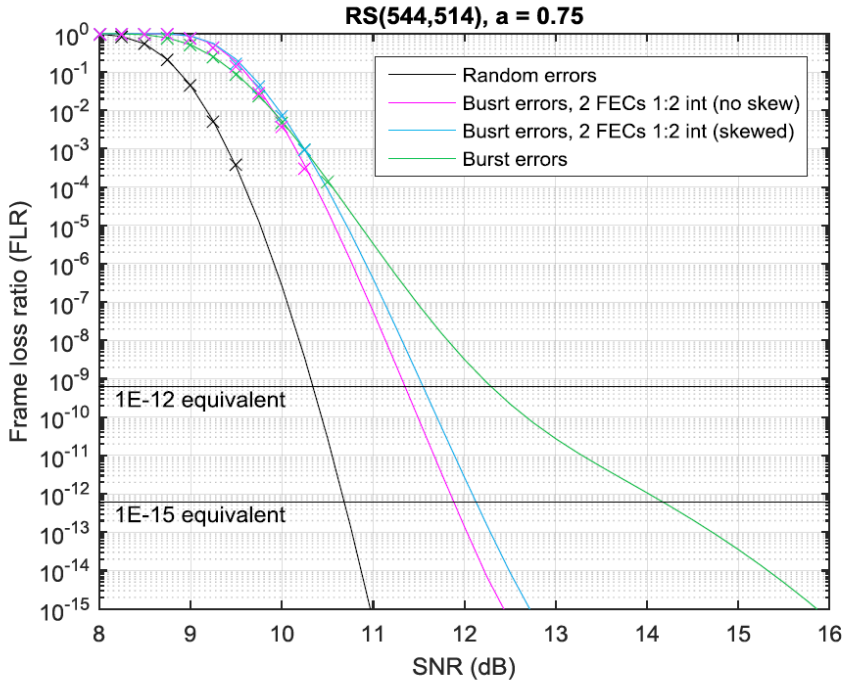

Figure 7 - RS(544) Single Part Link Performance (200GE)

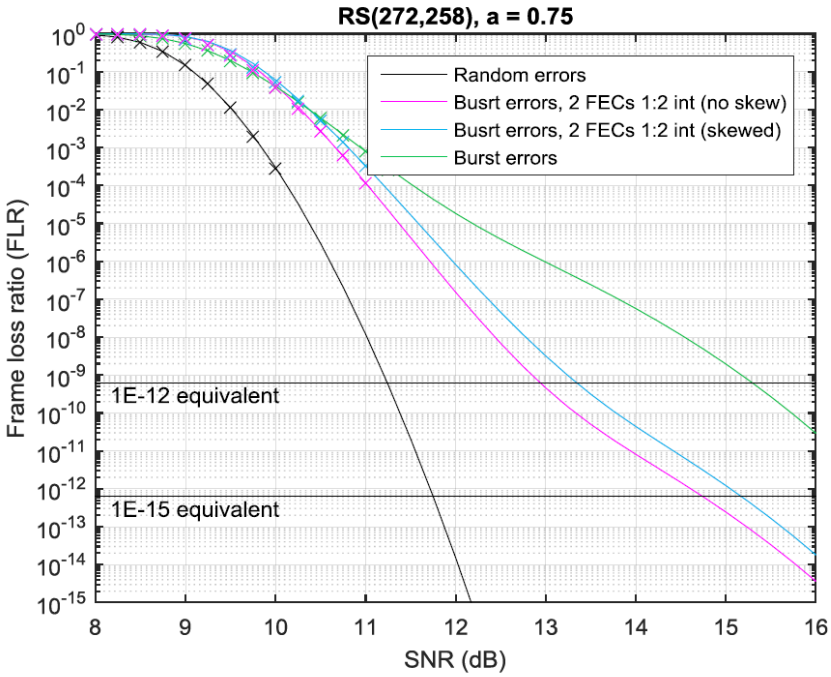

Figure 8 - RS(272) Single Part Link Performance (50, 100GE)

Table 2 - 200GE FEC Performance for RS(544) and RS(272)

FLR = 6.2E-10	RS(544,514)		RS(272,258)	
	BER	SNR (dB)	BER	SNR (dB)
Random errors	3.7677e-04	10.3438	9.9248e-05	11.2402
Burst errors, a=0.75	5.7878e-05	12.2968	8.8834e-09	15.2969
Burst errors w/ 2FEC 1:2 int (no skew), a=0.75	3.3131e-04	11.3496	1.3823e-05	12.9375
Burst errors w/ 2FEC 1:2 int (skewed), a=0.75	2.3646e-04	11.5469	5.0481e-06	13.3438

3.7 Annex B [Informative]: Example RS(272) Codewords

This annex provides example codewords produced by the PCS and FEC processes operating on a stream of idle control blocks, following the nomenclature of 802.3 Annex 91A (50 and 100GE) and 119A (200GE). The data is presented in hexadecimal form, transmitted to the symbol distribution function from left to right within each row starting from the top row and ending at the bottom row. The most significant bit of each hex symbol is transmitted first. In these examples the alignment markers are inserted as the first blocks of the codeword.

The scrambler seed used for the 200GE example is identical to 802.3 Annex 119A (S<0:57> = 24e6959d0fa5dbd), and the seed used for the 50 and 100GE examples is (S<0:57> = 0). For the 200GE example the value of the PRBS9 alignment marker pad seed was set to 0x00F immediately prior to am_mapped insertion. The additional 10 bit FEC pad symbol added immediately prior to the parity was set to 0 for simplicity in these examples. It should be noted that the 100GE example is associated with a two lane PMD using the appropriate alignment markers specified in 802.3cd (91.5.2.6).

Table 3 - Example tx_scrambled with alignment marker group for 50GBASE-R

Index <i:j>	**tx_scrambled_am<i;j>**
<0:256>	01284b775dc401007b1ed22488efbbff47a3d333db25e0028178b53093287ffff
<257:513>	10700000000f0001e7801e00000f0045e7879e00880f1e45e7779e1e88097e45e
<514:770>	1777955e89917e536776755ec291783367987540c2913c3360187248c28f3c3c9
<771:1027>	09098c48daef3c3c9056cc4ec2efa2bc8fb6ccf5c2f7c2ba01b6d2f5c9f3c3258
<1028:1284>	17175af437f387d586395a0257ed87d58d795b9a507b07d46d76109bb61b193a6
<1285:1541>	18d810dff21c9ebae2600ec1f2e36ea51c6676a1f2e365aa5df86eb7ec629a4a1
<1542:1798>	036368cf8c8a1454084327cb8a95934d805cd9c43e646ad398cab9da3e2b21958
<1799:2055>	121220cca1cac121ca4ba6aa3124df3fc10be33236b5d8362115a73dc13426c63
<2056:2312>	1b9ba73dc17b66c026eb39bdde94667ac6f347bb5094783ecdf0c6245885ee3f3
<2313:2569>	103082ca5f1d8ec9531e898a55c58f47d489e98bfecac2263207f77bfed486627

Table 4 – Example RS(272, 258) codeword for 50GBASE-R

Index <i:j>	**Codeword <i;j>**
<2719:2400>	a4290775dd010046f1bc22489fbbee17e2f65de626c3da034068c6564b0affcff0e0000030000f80
<2399:2080>	1e780004000f220e79e107910f02787a2ee79d78013e9a25d73df5522d134fcd9ddd5c6f5283d198
<2079:1760>	3cdc305686911870d83049c62e287927825c6448d6f3df0c24da88c0dfdd5144fb7ccdebfd00f975
<1759:1440>	602dbbd24ef3d3046b47ad6177f70ed5e30d4e02f50db5f18d4f4ed52ef0f071575843ec836c6ee4
<1439:1120>	1b01bb08ff3857975c647038374f6ca5638e64564fac75a7a51f9767ec6359452d8ad8e62630a145
<1119:800>	205847ca3a529933674036447cfac496a33aa773f89a830a35210114cce10d61294d7455a31c933e
<799:480>	ff4231f13b586e1b22196af320d0b8daf1ece721def4b3601bb0e6decbd316f3b1967ef05614a0f
<479:160>	9bc7d3191208fbd7e00e10d34fa5b8370a978d91a523af1be212979d1b7f352444c7e0eeffd2b661
<159:0>	e44004a57a215f04f00cb3ba55bd1f1c11fa557d

Table 5 - Example tx_scrambled with alignment marker group for 100GBASE-R

Index <i:j>	tx_scrambled_am<i;j>
<0:256>	10641906418b42d0b42d080200802003e0f83e0f9d2f4bd2f4bdff7fdff7ff5fd
<257:513>	1bd6fdfe0ca152a8a4043190660050100a008074d5ab15f6fef19b667fcff3fef
<514:770>	1f0a4b4fe6ec2367bb4d3540180201007d7d2a06c4b736112db28fe1fb7ebfe75
<771:1027>	158f07ac9d47ee35b2343d04c3300c42903c1458be087393b2d0bacd33fcbfbff
<1028:1284>	1fe071ab54c986197db2a74e0180200006e39522cf9e79a093502cffdffff9feb
<1285:1541>	10700000000f0001e7801e00000f0045e7879e00880f1e45e7779e1e88097e45e
<1542:1798>	1777955e89917e536776755ec291783367987540c2913c3360187248c28f3c3c9
<1799:2055>	09098c48daef3c3c9056cc4ec2efa2bc8fb6ccf5c2f7c2ba01b6d2f5c9f3c3258
<2056:2312>	17175af437f387d586395a0257ed87d58d795b9a507b07d46d76109bb61b193a6
<2313:2569>	18d810dff21c9ebae2600ec1f2e36ea51c6676a1f2e365aa5df86eb7ec629a4a1

Table 6 – Example RS(272, 258) codeword for 100GBASE-R

Index <i:j>	Codeword <i;j>
<2719:2400>	304c1304c11685a1685a00802008023e0f83e0f8a5e97a5e97fdff7fdff7d7ddf6bfbf07c534a915
<2399:2080>	24c209866a008050100e02b2d5ba8f6cf7d9a663fcff7fdff4a25acfc6ed8bbc659590040300801
<2079:1760>	d7e97c0a469d90d6929bfe3f0fdbfa5cc6b83cd7ae5f8b1d360b02f0c8338c250f00a04681f38727
<1759:1440>	d3742cd7324ff7fffc3fc716a9970c4c2df2a797c020000300e3a549a3cfcf082565a0ffbfdfffc
<1439:1120>	e80e00000f00007801e780000000fa21e79e007110f027a7aeec79179017eba2dd53df5722d114f
<1119:800>	cd8dd55c6f1283dd9b3c5c205287919840dc3249861e287927424c6c49d6f3cf0824dac8d0d7df51
<799:480>	c4db7cccebbd10f177602dabde4cf393346747ad6167ff0cd5e32d4203f50d95f58d4fcec526f1f0
<479:160>	b1775842ecc34c62e71b81bb04fd38d787506770f807476ca563ae68574f6c55aba61fd777e46259
<159:0>	8500014c5e402e3ed2dfcf329fd78d06a09c7a96

Table 7 - Example tx_scrambled with alignment marker group for 200GBASE-R

<0:256>	0b2acab2acab2acab2aca9324c9324c9324c9324c942108c22c90a2697a225314
<257:513>	1aa6a9ba6e98a629ab6adab6adab6adab6ad9be6d9b26ddb76e5b46def365bdf1
<514:770>	0c42a493a18d4f681a02679e4456dd1c98cfc35417a7b1742db5334a1bd2041b7
<771:1027>	07ace48b2fd61333bfb75c3dc524204cebd0955d1b04967a00f7c5cc8253b479c
<1028:1284>	007000000000e0001b8001c0000f000567806e00090f0185e7899e01480f7245e
<1285:1541>	019f3d6d103a1c8aba6f2f7d1bf3fcf0667c38bf1c33c468e6b6484fdd3c76a93
<1542:1798>	0420d6343af3153567c06b24d60acd8fbe36dcae9dcf0e56d70cb5bbde6d8cdeb
<1799:2055>	09536bdc4c14c98debc2de801ee5e6c7f70a6bca3697a52a5177ddc7f14ba52fd
<2056:2312>	045bcd6c1a9a2073c9466f46411a902cda99e4b01f1138ffc1c70e763bcdb1ec9
<2313:2569>	1b180dda54aec655752a2b19815bc10b7e0871e2400c523c411c60347b96b7d02
<2570:2826>	1fae40e28be4b4962907f1cbd57f2a69d1d68ca18c69376d313bfa0dc02facb85
<2827:3083>	0a8db852f683475a0103123314d3d0a4216d2b647cb6f35e9433b4f99094b0ccf
<3084:3340>	1fe82e0b79a7b623eb78cdb22e3b7961bdef1a794da5865b8eddd55fde7c2d490
<3341:3597>	11e35f5a1a35623acb63bc42f8ce49ca5a9e8c2da7274f8efb6d45f6d42539353
<3598:3854>	0d826f8ec6479e447a976c8d1912bfc7b8690639cbdf23e26e563035331454fbf
<3855:4111>	1fa44bca65d535687a5958f3a5b9e8af769d4f3830971143d0b2e02ea26535347
<4112:4368>	03898385e35d893fee265a6aa8be034f5b47e6acb1365c1b8f1bc7947a89316ce
<4369:4625>	1c79eef8f72045aae815feb69dc141476fc42f25fe8fd985bd76d6cca88c8cc2c
<4626:4882>	09b954674d5513caa74fb3f4c00a6dce3dece9eb9ee0aa5cab87bbb3017d985d8
<4883:5139>	0fa82fe1c8d37e9c2e1a5e8f0c8393b69f0ea4e44ddaf0ee0cb28cf2aaa55e9ac
<0:256>	0b2acab2acab2acab2aca9324c9324c9324c9324c942108c22c90a2697a225314

Table 8 – Example RS(272,258) codeword A for 200GBASE-R

Index <i:j>	**cx_A<i;j>**
<2719:2400>	a6a9aa6a9a64992649921486284bd26519565d946d5b56d5b57d936ed8b6cf7efa842ede6024f1db
<2399:2080>	63156bc9a159a5edb273fd0ccbb63b810bd2ea92c033a728390000000000000000000e00c063c401de
<2079:1760>	83daf5c354f4ff60f30f38d886d6fc6e21358c67cd66ba0ef876ddcd4e98f79982a63b4cbd705fa7
<1759:1440>	bfbd6a5a54efa8fd2bb4c1808148c54aeb6691111fce19ee6f018a9a615219a0f07d1ea3208c05a7
<1439:1120>	40813fa349fc15ecab5c30b6491982be950874161688cca2502dbe2b3cc267cd2ffe0ef2ef8d9b8e
<1119:800>	c371ed974cddb7f5a1e25adabd348f732943153971cad2b727835b8f3caf626a43be30f7b2318c66
<799:480>	be625996ac695cf5e56e73ce92f0079922c0c98fe48c8aa98016e6ad35e34f122b99de4ed56faae5
<479:160>	2808ffa66febccd1308dc5555aa66f0079de6fce4ab0e198cd7c07fb2f435e8135b657ecb7053554
<159:0>	5e8001d646735b2e0f58cf1ec4d6e608792c4af7

Table 9 – Example RS(272,258) codeword B for 200GBASE-R

Index <i:j>	**cx_B<i;j>**
<2719:2400>	a6a9aa6a9a6499264992109a2c8a226519565d946d5b56d5b56dbb69dbb66986849163c0bcc51271
<2399:2080>	1fbd07455bc28405e74486bf7c3892732a40d9793ed04e2c70000383b01c1e3350ef84f43cc02512
<2079:1760>	3e48b4e3652fbfc99bd1f327184bcb25560ae19501d64d998fbaa1e75bb5db1afd6e0cb1ad0780d9
<1759:1440>	50f14a5fa2e39744bf59b29e7ecc41683ccf038ec72371b4dc6ef75ba8d46a2d038402f123878bed
<1439:1120>	afe8e691289c5fd5ca62b196dfec07e9ac5f4b8b08331bca106d5b4978b7908cc82cf68d8f6898f6
<1119:800>	f78f234b8e555f3127ac161c4719f4e4b956d32eb7efa292cafb04c112dd4c47f4b273e254e56145
<799:480>	ffcf4abb851ab69f50ae0c0a24d15dac8e3438ddeff2dfa179cfe46b05ecbc74678f1f88017b5e83
<479:160>	db8f4c5da1ad9150cd3b973919725fb2b78bca0ea9fbbf43744144e97d610f3270f89cf58c1f3295
<159:0>	35800311f0cf1bcdc405de6316348b9f01e49c4f