Two ways of showing 25G-AUI compliance

Piers Dawe
Mellanox Technologies

Supporters

- Adee Ran — Intel
- Tom Palkert — Luxtera
- Mike Dudek — QLogic

Introduction

- This presentation summarises the new material in Annex 109B
  - Which was introduced following "FEC-protected chip-to-module 25G-AUI specification"
    - dawe_3by_01_0315.pdf and dawe_3by_01a_0315.pdf
  - D0.1 comment 145 "Implement with editorial license the modifications that apply to the module interface in dawe_3by_01a_0315 Slides 10 and 11.
- There are two ways for a module to show 25G-AUI compliance
  - What clean-up and bug fixing is needed in the draft?
  - Do we want to allow the host a similar choice of two ways?

Major options, motivation

- D1.0 has 3 "major options"
  1. Does PHY* include^ an RS-FEC sublayer?
     - If so, can qualify the module to the C2M CAUI-4 procedures and specs, or the "alternative" Annex 109B procedures and specs
       - 2, 3 Two sub-options, for module output eye and for module stressed input performance. Other specs are the same as C2M CAUI-4
     - If not^, qualify the module to the C2M CAUI-4 procedures and specs
- The Annex 109B procedures and specs are less onerous
- The CAUI-4 procedures and specs would be required anyway for a module that also supports 100GBASE-SR4 and C2M CAUI-4
- The host doesn't care – all module options work

* That's this PHY using 25G-AUI, not the same silicon in 25GBASE-CR mode

^ There are no PHYs that use 25G-AUI that don't use RS-FEC, unless you put the FEC in the module, which is unlikely for SFP or QSFP

Module stressed input procedures and specs

- Module's input must run at a specified BER when presented with a stressed eye with specified eye height and eye width
- For C2M CAUI-4, these dimensions are specified at the **1e-15** points, found by extrapolating bathtub curves. The module input's BER must be better than **1e-15**
  - It could be difficult to show this, as a back-to-back 25GBASE-SR link is allowed to make some errors, that RS-FEC corrects
  - Measuring 1e-15 is very slow: 10 hours per error
- For the Annex 109B way, the dimensions are specified at the **1e-8** points. Same height and width numbers. The module input's BER must be better than **1e-6**
  - Very fast measurement for the module
  - This would allow a relaxed spec for the host, which isn't in D1.0
  - 1e-6 is << 5.2e-5 total of 25G-AUI and 25GBASE-SR errors, which RS-FEC can correct to 1e-12, the objective spec

Module output eye

- Module output eye is defined by (inner) eye height and width at a certain percentile, and VEC which is the ratio of eye amplitude to eye height
- Found by drawing bathtub curves and extrapolating. No errors are counted at the headline BER
- For C2M CAUI-4, specified at the **1e-15** points, found by extrapolating bathtub curves from **1e-4** to **1e-6**
  - Needs an effective 4 million samples. Time-consuming to do properly with a sampling scope
- For the Annex 109B way, specified at the **1e-8** points, found by extrapolating bathtub curves from **1e-3** to **1e-5**. Same height and width numbers
  - Needs an effective 400,000 samples. 10x faster to do properly
    - Slightly more pessimistic than extrapolating from 1e-4 - 1e-6, but not extrapolating far, and reduced test time is more important
  - If anyone wants to measure both ways, it's the same measurement with different curve fitting, so no extra time needed

Interoperability with the host?

- Host to module
  - Host delivers 1e-15 at particular dimensions
  - Module input's errors are either:
    - <1e-15 with this eye (CAUI-4 way), or
    - <1e-6 BER with a 1e-8 eye of the same size
  - Either way delivers better than 1e-6
- Module to host
  - Module delivers 1e-15 or 1e-8 at particular dimensions
  - Host is tested to 1e-15 BER for the same dimensions
  - Host can be expected to deliver better than 1e-6, perhaps 1e-8, with the eye the module is allowed to generate
    - Using a mix of 1e-6 and 1e-8 to build in more conservatism here, to avoid any extra host specs for a dual-use CAUI-4 / 25G-AUI host

Should we give the host the same options?

- The methodology supports it
- Hosts that support 100GBASE-LR4 would not benefit
  - But don't have to take the options
- Hosts that don't, such as 25G NICs in servers, would benefit
  - Again, it doesn't matter if the NIC supports 25GBASE-CR without RS-FEC because that's not 25G-AUI

Changes to give the host the same options

- In 109B.3.1, (25G-AUI C2M host output characteristics, change:
- A 25G-AUI C2M host output shall meet all specifications in 83E.3.1.
  - to
- A 25G-AUI C2M host output shall meet all specifications in 83E.3.1 with the exception of eye height and eye width. A 25G-AUI C2M host output also shall meet the eye height and eye width specified in 109B.3.1.1.
- Insert new
- 109B.3.1.1 25G-AUI C2M host output eye opening
- with contents based on 83E.3.1.6, in the style of 109B.3.2.1, with the same differences that 109B.3.2.1 has in relation to 83E.3.2.1. Note there is no vertical eye closure spec for host output
- In 109B.3.3, 25G-AUI C2M host input characteristics, change:
- A 25G-AUI C2M host input shall meet all specifications in 83E.3.3.
  - to
- The 25G-AUI C2M host input shall meet all specifications in 83E.3.3, with the exception of the host stressed input test.
- For a PHY that does not include an RS-FEC sublayer (Clause 108), the 25G-AUI C2M host input shall meet the host stressed input test requirements in 83E.3.3.2.
- For a PHY that includes an RS-FEC sublayer, the 25G-AUI C2M host input shall meet the host stressed input test requirements in either 83E.3.3.2 or 109B.3.3.1.
- Insert new
- 109B.3.3.1 25G-AUI C2M host stressed input test
- with contents based on 83E.3.3.2, in the style of 109B.3.4.1, with the same differences that 109B.3.4.1 has in relation to 83E.3.4.1, with a VEC6 in the range of 3.5 dB to 4.5 dB with a target value of 4 dB. VEC6 is defined as 20*log10(AV/EH6)
- Apply the changes to 109B.3 and 109B.4 agreed for other comments to the new material
- Revise the PICS to follow these changes