Adaptive vs. Programmable CTLE RX:

Cisco feedback

April 24, 2014

Content

This presentation share our point of view about mode of operation of the module's input RX CTLE, as currently defined into IEEEdraft 802.3bm/D2.2 and OIF_CEI_03.1.

It provides concerns and proposed path to lower risk development for future hosts and modules supporting RX CTLE functionalities.

Supporters

Ali Ghiasi - Ghiasi Quantum LLC
Andre Szczepanek - Inphi
Angus Li - Macom
Atul Gupta - Macom
David Brown – Semtech
Edward Frlan – Semtech
Francois Tremblay - Semtech
Henry Chen - Broadcom
Howard Frazier - Broadcom
Jeffery Maki – Juniper
Justin Abbott - Semtech
Marek Tlalka - Macom
Mark Gustlin - Xilinx
Mike Harwood - Inphi
Mike Furlong - Clariphy
Mike Peng Li - Altera
Neel Patel - Clariphy
Norm Swenson - Clariphy
Ryan Latchmann - Macom
Scott Kipp - Brocade
Vivek Telang – Broadcom

Editors and contributors

Alessandro Cavaciuti - Cisco
Brian Eldridge - Cisco
Gary Nicholl - Cisco
Gianpiero Bognanni - Cisco
Marco Mazzini - Cisco
Pirooz Tooyserkani - Cisco
Vineet Salunke – Cisco

Thanks to:
Ali Ghiasi - Ghiasi Quantum LLC
for slide 12 contribution

Background

Starting from CEI-28G-VSR revision 12 (Sept 2013), the host is required to provide to the module RX input a recommended CTLE peaking value.
This value should be defined during host validation at TP1a.

The recommended value should be one out of the nine defined into Table 83E-2 and Figure 83E-10.

This methodology has been promptly adopted into OIF_CEI_03.1 and IEEEdraft 802.3bm/D2.2, requiring the host to write this value into the programmable module via a dedicated variable «*Recommended_CTLE_value*».

Table 83E–2—Reference CTLE coefficients

Peaking (dB)	G	$\frac{P_1}{2\pi}$	$\frac{P_2}{2\pi}$	$\frac{Z_1}{2\pi}$
1	0.89125	18.6	14.1	8.364
2	0.79433	18.6	14.1	7.099
3	0.70795	15.6	14.1	5.676
4	0.63096	15.6	14.1	4.9601
5	0.56234	15.6	14.1	4.358
6	0.50119	15.6	14.1	3.844
7	0.44668	15.6	14.1	3.399
8	0.39811	15.6	14.1	3.012
9	0.35481	15.6	14.1	2.672

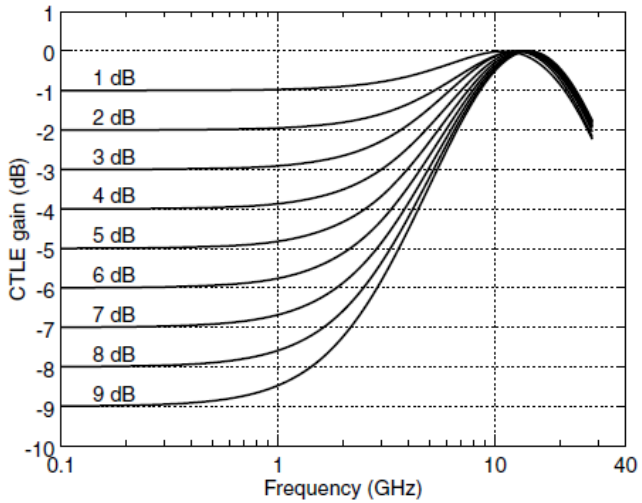


Figure 83E–10—Selectable continuous time linear equalizer (CTLE) characteristic

We're concerned to adopt this methodology, since we believe that:

1. It's not reflecting industry's status of the art, not addressing what we asked for and not targetting for future standards (400G).

2. Opens future development to marginal designs, which may guarantee minimal performances over IEEE compliance conditions, thus to potential technical issues
   - Potential failure case described into next slides
   - Potential non-symmetrical IEEE/OIF host/module and module/host channels.

3. Will add more costs due to testing time, for both hosts and modules manufacturers.

We are going to provide arguments for above three points in the next slides.

1. It’s not reflecting industry’s status of the art

Our point of view

- 100G hosts, based on 25G electrical lanes, have been under development since 2010, and Cisco, as host manufacturer, has been one of the active players in the definition of CEI-28G-VSR budget and calibration rules.

- As a matter of fact, all CEI-28G-VSR compliant phys that have been evaluated by us between this period supported full-adaptation or plan to add it, for both host and module purposes (even if designed for module application, most of the parts have also TX FIR block that allow to deliver opened eyes at TP1a (CAUI-4 CTLE=1) to the module).

- Recently are also available some parts with some limitations w/respect the continuous adaptive (but still better than a programmable module method):
  - Single CTLE RX setting ensuring compliance over different channel losses (correnspondingto different CTLE coefficients).
  - Adaptive at start-up mode, that allow the CTLE RX to optimize for a given channel/data rate (eye monitoring, BER measurement, etc..).

1. It’s not addressing what we asked for and it’s not targetting for future standards (400G).

- Although never quoted directly into CEI-28G-VSR draft 00 (10/26/2010) to 11 (07/25/2013), the common way to think of the host/module CTLE receivers was a full-adaptive mode to the module/host input conditions over the VSR channel.

- In July 2012, we asked to add in the OIF draft that the VSR receiver should be continuously adaptive to cope with ageing and V/T variations. Our comment to OIF was accepted (oif2012.253.02.xlsx, row 160), but never reported in any VSR drafts.
  - “Add appendix 1.C titled 'VSR receiver': Content 'The specification assumes that the receiver equalization function is continuously adaptive'. The receiver should equalize channel variations including those due to temperature changes as well as transmitter output variations due to temperature and voltage”.

- We think that a programmable module implementation is not the only allowable solution moving forward for IEEE 802.3bm, which was drafted more than 3 years late than OIF-28G-VSR. We think that the same method is not providing any smart, cheaper and of easier implementation for future 400G hosts and interfaces

2. Opens future development of marginal parts, thus to potential technical issues

Technical concerns hypothesis: 83E.3.1.6 Host output eye width and eye height

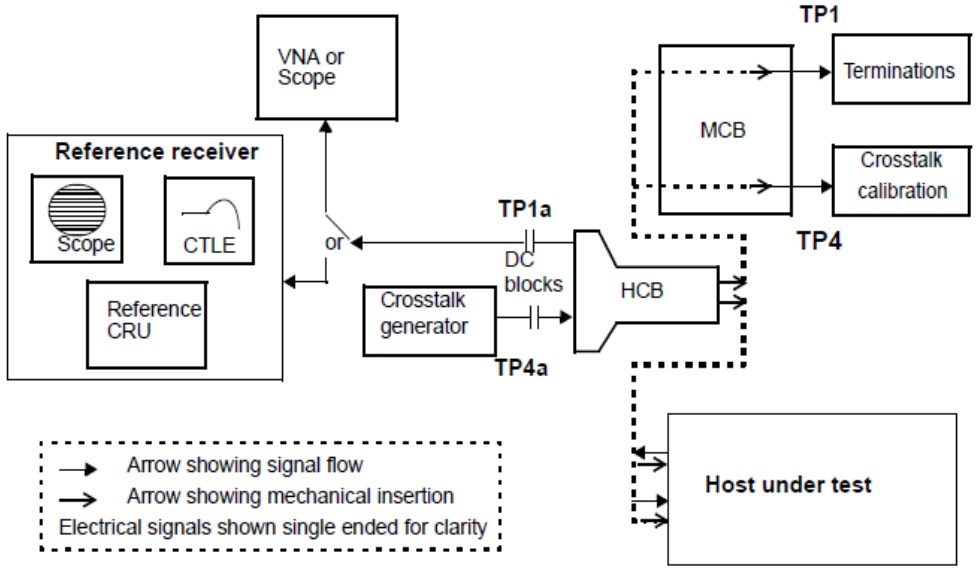


Figure 83E–9—Example host output eye width and eye height test configuration

83E.3.1.6.1 Reference receiver for host output eye width and eye height evaluation

The reference receiver is used to measure host eye width and eye height. The reference receiver includes a selectable continuous time linear equalizer (CTLE) which is described by Equation (83E–4) with coefficients given in Table 83E–2 and illustrated in Figure 83E–10. The equalizer may be implemented in software, however the measured signal is not averaged.

Apply the reference receiver including the appropriate CTLE to the captured signal. For modules, any single CTLE setting as described in 83E.3.2.1.1 which meets both eye width and eye height requirements is acceptable. For host compliance, the CTLE peaking in the reference receiver shall be set to one of three values. These are: a) the recommended CTLE peaking value provided by the host, b) the value 1 dB higher if present in Table 83E–2, c) the value 1 dB lower if present in Table 83E–2. Any of the three CTLE settings that meets both eye width and eye height defined in Table 83E–1 is acceptable.

From 83.E.4.2 - Technical concern built assuming three equalizer's settings allowing TP1a compliance.

2. Opens future development of marginal parts, thus to potential technical issues

Technical concerns hypothesis: 83E.4.2.1 – Module stressed input test

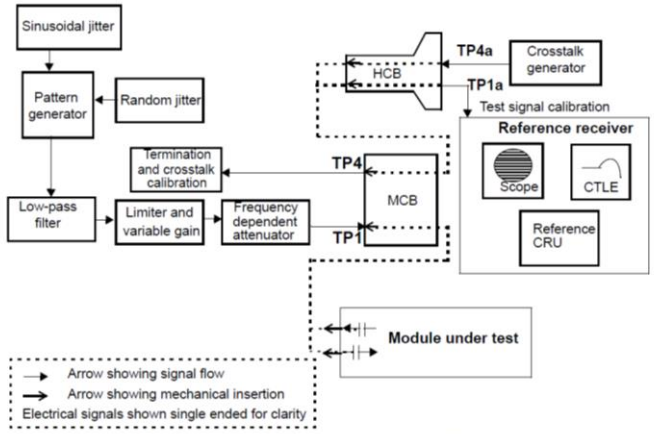


Figure 83E–15—Example module stressed input test

Two levels of frequency dependent attenuation are used for the module stressed input test: high loss, and low loss. For the high loss case, frequency dependent attenuation is added such that from the output of the pattern generator to TP1a is 10.25 dB loss at 12.89 GHz. Eye height and eye width are then measured at TP1a using the setting of the software CTLE which maximizes the product of eye height and eye width based on the eye measurement methodology given in 83E.4.2. Random jitter and the pattern generator output amplitude are adjusted to result in the eye height and eye width given in Table 83E–8 using the reference receiver. For the low loss case, discrete frequency dependent attenuation is removed such that from the output of the pattern generator to TP1a comprises the mated HCB/MCB pair as described in 83E.4.1. Eye height and eye width are then measured at TP1a using the setting of the software CTLE which maximizes the product of eye height and eye width based on the eye measurement methodology given in 83E.4.2. Random jitter and the pattern generator output amplitude are adjusted to result in the eye height and eye width given in Table 83E–8 using the reference receiver. In both the low loss and high loss cases, the module under test is provided with the reference CTLE setting used to meet eye width and eye height requirements via the variable *Recommended_CTLE_value*. If a Clause 45 MDIO is implemented, this variable is accessible through register 1.169 (see 45.2.1.92a). The module under test is evaluated with three *Recommended_CTLE_value* values for both the high loss test and low loss test. These are: a) the CTLE setting used to meet eye width and eye height requirements, b) the value 1 dB higher if present in Table 83E–2, c) the value 1 dB lower if present in Table 83E–2.

Should be clearly stated that the module has to meet EW and EH requirements at least for the Recommended_CTLE_value AND the +/-1dB cases ?

Note: this procedure is not needed for adaptive modules.

2. Opens future development of marginal parts, thus to potential technical issues

Technical concerns – channel ageing and V/T effects - other hypothesis -

- As said before, we asked to mention in the OIF standard text that the VSR receiver is continuously adaptive to cope with ageing and V/T variations.

- Ageing estimation were already given by other by simulation into OIF presentations. Cisco was always looking for a margin of 2dB on the VSR channel loss (as shared into oif2010.132.01.pdf) – for this we tested CDR and Gearboxes using a 12dB channel, to include ageing and other impairments.

- During the validation of RX CTLE parts having CTLE monitoring, it's normal to observe a variation of around +/- 1 dB on the input RX CTLE gain value over T/V corners (begin of life).

- TP1a host characterization is done leveraging a reference receiver which is typically implement with SW CTLE filter. This could lead to discrepancies with a real chip CTLE implementation, this reference receiver does not capture non-idealities and design tradeoffs built into the module receiver. Besides this we have to take into account the fact that TP1a and TP1 test are done in different ways, equipment tolerances, etc... If we assume optimum working point for host and module we have to assume that at least 1dB offset between host and module test on the RX CTLE coefficient can happen.

2. Technical Concerns if marginal parts will be adopted.

Potential link broken case when compliant host is mated to a compliant module.

Example – case of multiple points of compliance for host (TP1a) and module (TP1), under minimal IEEE compliance conditions.

For some port cases, the host is characterized and qualified to be TP1a compliant with multiple CTLE coefficients (e.g. 5,6 and 7dB, beyond IEEE request) – 6 is the optimum coefficient choose for IEEE implementation.

The module is also characterized and qualified by the module manufacturer to be IEEE TP1 compliant with CTLE RX coefficient setting to 6, 7 and 8 dB, with the others being not compliant (or not tested).

The host write 6 (optimum value) to the module -> module's fall into a marginal compliance state.

Host/module ageing and environmental changed conditions make the channel (host TX and PCB + connector + module's PCB and RX) to degrade by 1dB.

The channel is broken because the module's RX coefficient during time became 7dB, but the host will not write into the module anymore-> link is broken, should worked with adaptive mode.

Even in case of re-programmability, the host doesn't have any info about the fact that 7 and 8 RX CTLE coefficients can be also used by the module.

Neither host and module are responsible for the link broken due to ageing -> responsibility is shared.

Failures can happen even assuming multiple points of compliances (beyond current IEEE requirements) for the host.

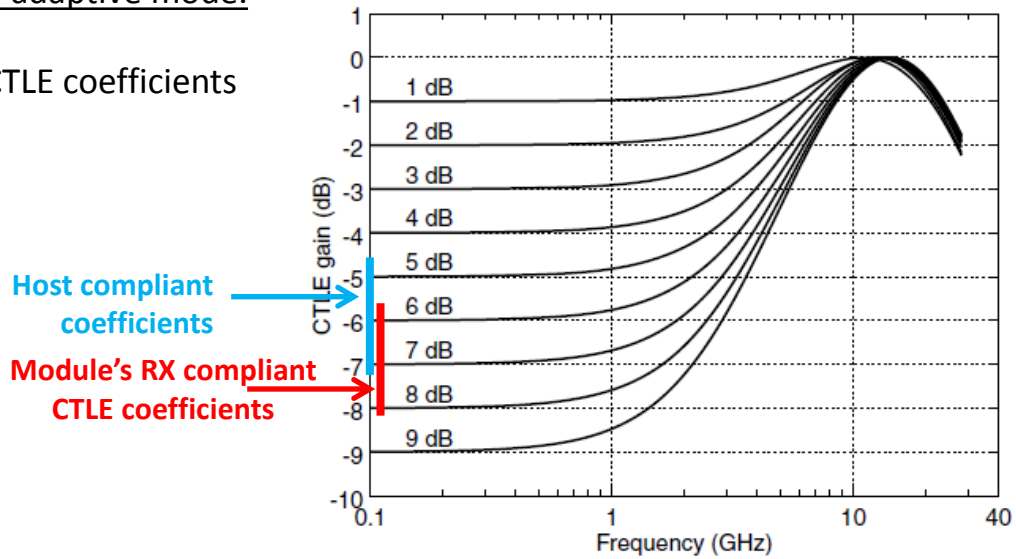
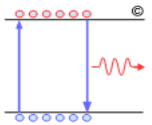


Figure 83E–10—Selectable continuous time linear equalizer (CTLE) characteristic

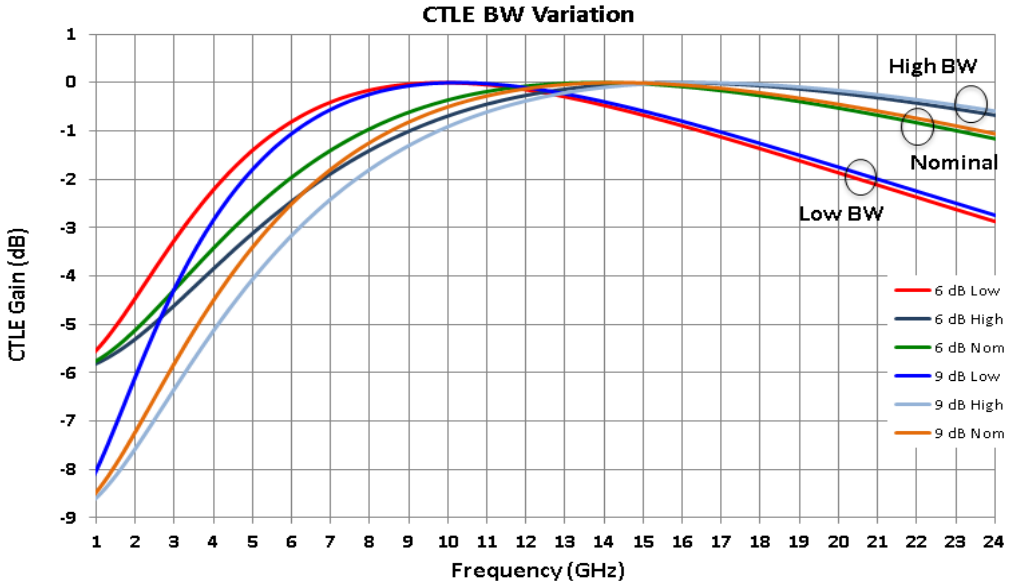
Non Adaptive CTLE Impacts the System Margin

- **Effective CTLE gain and how well it can equalize the channel can vary when CTLE is not adaptive due to:**
  - Set point accuracy
  - Loss and termination differences between measured point TP1a and actual module IC receiver
  - Loss, termination, and PCB shape differences between the actual host and module output plus some reference PCB
  - A lesser know variation is due to CTLE poles/zero BW variations
    - It is assumed poles/zero BW can go as low as the ratio of 10/14 and as high as the ration of 16/14

CTLE Variation Due to Poles/Zero Tolerance



- **For the 6 dB and 9 dB peaking**
  - CTLE gain vary by ~ 0.5 dB



2. Technical Concerns if marginal parts will be adopted

Potential non-symmetrical IEEE/OIF host/module and module/host channels.

Said that we already deployed hosts and modules potentially supporting up to 12dB loss (to include ageing and V/T effect) we are now concerned that:

- The introduction of marginal programmable RX CTLE modules might oblige us to reconsider our design rules, targeting narrower range of losses or more expensive materials even respect to the 10dB budget (because we believe that the current IEEE definition of compliance ranges for the programmable host/module can lead into potential case of broken links, as shown in slide 11).

- Or, since the Module TX to Host RX path will still be supported by continuous adaptive RX CTLE we might fall into designs constraint that will exacerbate channels asymmetricity, depending on the traffic flow (shorter channel from host to module, longer from module to host).

Both paths would be painful for us, especially for high density port design.

3. Will add more costs due to testing time on both host and modules.

To implement the capability of the host to write into the module, a port to port host characterization will have to be run to verify which is the CTLE optimum value.
This is already foreseen into IEEE, as well done over current hosts.

But we also have to take into account that:

- It is not granted that all VSR receiving lanes into the module will share same optimum CTLE set-point and passing range -> need module's detailed characterization.

- Host compliance is accomplished using a given host transmitter device while stressed receiver sensitivity uses external equipment which is inherently a different implementation than the host device. This can lead to non-optimal CTLE set-points within the module with respect the host ones.

- With programmable module RX CTLE gain is provisioned statically by the host, based on a specific calibration procedure, then the matching requirement of the CTLE implementation to the standard definition becomes more stringent for a given device. No need to care about how well module device CTLE implementation matches CEI-28G-VSR/CAUI-4 CTLE curves for adaptive modules, since algorithm will adjust and settle to the per lane optimum CTLE setting.

3. Will add more costs due to testing time for both host and modules manufacturers.

Giving the previous shared technical concerns and points, we believe that it will be more complex to qualify a programmable RX CTLE (chip or whole module).

- More testing time for HW verification (test that each single RX CTLE setting works, since the programmability does not allow to adjust to another better gain), considering potential offsets.
  - The programmable module should include a look up table of the complaint RX_CTLE coefficient range in its EEPROM (R), for both short and long channel cases defined by IEEE. This would help to convert to the characteristics of the actual CDR used and allow further settings from the host if needed (similar activity agreed with some module manufacturers on SFP+ development – all these info are already available from foreseen TP1 module stressed input test).
  - Suppliers would have to provide evidence that their chip / module can work over a wider range with respect host settings +/-1dB, to cope with ageing and environmental variations.

- For this there'll the need of an early engagement between chip/module suppliers and host manufacturers, starting during development phase.
  - To agree on common tests for good correlation of host and module CTLE coefficients and manage coefficient tolerances.
  - To define parts that can be used into modules, based on stand-alone chip characterizations.

- After module's qualification, we expect more testing time for SW implementation and settings verification during integration phase.

Comments

- We don't believe that programmable RX CTLE methodology as currently defined into OIF_CEI_03.1 and IEEEdraft 802.3bm/D2.2 is the right path, since we believe that:

1. It's not reflecting industry's status of the art, not addressing what we asked for and not targetting for future standards (400G).

2. Opens future development of marginal designs which may guarantee minimal performances over IEEE compliance conditions, thus to potential technical issues

3. Will add more costs due to testing time, for both hosts and modules manufacturers.

4. Does not provide any clear positive advantage.
   - Power consumption? Phy suppliers are encouraged to provide power numbers comparing adaptive and programmable modes.
   - Lower costs of settable parts? Module/phy suppliers are encouraged to provide % of cost savings.

- For this currently we plan not to consider devices implementing only programmable CTLE RX.

Proposed path

Unless it would be strictly needed, parts supporting adaptation will have priority in our developments.

However, we understand people which are looking to define programmable RX CTLE parts (already in the field), and their need to define viable method to keep these in this market.

For this we propose to add a definition for adaptive solutions as a preferred choice into standards and MSAs, and also keep the programmable mode of operation.

Proposed path for IEEE and OIF

We would like to have:

- Updated both IEEE and OIF specifications to make it clear that both adaptive and programmable CTLE options are acceptable (it isn't right now, only programmable mode is clearly defined).
  - Add "Adaptive CTLE" support as a line item to 83E.5.4.4 (Module Input) PICS of type "O" (optional).
  - Add an ability bit for "Adaptive CTLE" into Clause 45, preferably adjacent to the recommended peaking register location at 1.179.
  - Proposed wording for the adaptive CTLE: " The adaptive module shall autonomously determine an initial CTLE gain setting immediately upon start-up . After start-up the module shall enter into a slow continuously adaptive mode, such that it is able track channel variations ".
- Update in the 802.3bm standard to capture and address any differences in compliance testing between the two options, addressing our technical concerns about TP1 compliance range for programmable parts.

Proposed path for MSAs

From MSA we would like to have:

- A clear indication in module's management interface whether a module is using adaptive or programmable equalization (e.g. adding a bit in Upper Memory Page 00h, Address 193 of SFF-8636) - this should be clearly available to the host, reading into module's register.
  - CDR_RESET register has to be available for both module's TX and RX CDR (if not yet defined) and should be on per-lane basis.
  - Define and include a look up table of the compliant range to TP1 stress test RX_CTLE coefficient into the module's EEPROM (readable by the host), for short and long (10.25dB) channels.

In next slide Cisco provides a proposal for Module Bring Up Sequence related to module's RX CTLE adaptation.

Module Bring Up Sequence related to CTLE adaptation – proposal to MSA

1. Module is powered up and held in reset.
2. Reset is removed from the module and IDPROM is read.
3. Power_level is set in the module appropriately.
4. Transmitter from the Host Phy to module is turned on (adjust VMA, etc..).
5. MAC/PCS is brought up -> idles are being transmitted towards the module.
6. Trigger the adapting case when writing the recommended RX CTLE module's setting from the host (per individual lanes, here the TX CDR start-up begins, getting valid data).
7. At this time the CDR in the module should start adapting to the incoming data (< 100ms).
8. TX is enabled in the module (turn on laser).

THANK YOU

Back-up

IEEE 802.3bm and OIF CEI-28G-VSR Implementation Agreement - Draft Revisions History

10/30/2013

CEI-28G-VSR Implementation Agreement Draft 15.0 Document: OIF2010.404.15

13.3.2 Host-to-Module Electrical Specifications

Each host-to-module lane shall meet the specifications of Table 13-1 and Table 13-2. Definitions and methodologies can be found in Sections 1.3.4 to 1.3.11. The host shall provide a recommended CTLE peaking value selected from Table 13-8 such that the requirements defined in section 13.3.11.1.1 are met. The method of providing this is outside the scope of this document.

09/18/2013
09/17/2013
09/16/2013

CEI-28G-VSR Implementation Agreement Draft 14.0 Document: OIF2010.404.14

CEI-28G-VSR Implementation Agreement Draft 13.0 Document: OIF2010.404.13

CEI-28G-VSR Implementation Agreement Draft 12.0 Document: OIF2010.404.12

13.3.2 Host-to-Module Electrical Specifications

Each host-to-module lane shall meet the specifications of Table 13-1 and Table 13-2.
Definitions and methodologies can be found in Sections 1.3.4 to 1.3.11. The host shall provide a recommended CTLE peaking value selected from Table 13-8. The method of providing this is outside the scope of this document.

Table 13-1. Host-to-Module Electrical Specifications at TP1a (host output)

07/25/2013

CEI-28G-VSR Implementation Agreement Draft 11.0 Document: OIF2010.404.11

13.3.2 Host-to-Module Electrical Specifications

Each host-to-module lane shall meet the specifications of Table 13-1 and Table 13-2. Definitions and methodologies can be found in Sections 1.3.4 to 1.3.11

Table 13-1. Host-to-Module Electrical Specifications at TP1a (host output)

Draft Amendment to IEEE Std 802.3-2012
IEEE P802.3bm 40 Gb/s and 100 Gb/s Fiber Optic Task Force

IEEE *Draft* P802.3bm/D2.2
5th April 2014

Draft Amendment to IEEE Std 802.3-2012
IEEE P802.3bm 40 Gb/s and 100 Gb/s Fiber Optic Task Force

IEEE *Draft* P802.3bm/D2.1
7th February 2014

83E.3.1.5 Transition time

The transition times (rise and fall times) are defined in 86A.5.3.3.

83E.3.1.6 Host output eye width and eye height

Host output eye width is greater than 0.46 UI. Host output eye height is greater than 95 mV. Figure 83E–9 depicts an example host output eye width and eye height test configuration. Host output eye width and eye height are measured at TP1a using compliance boards defined in 83E.2. The host output eye is measured using a reference receiver with a continuous time linear equalizer (CTLE) defined in 83E.3.1.6.1. The recommended CTLE peaking value is provided to the module via the variable *Recommended_CTLE_value*. If a Clause 45 MDIO is implemented, this variable is accessible through register 1.169 (see 45.2.1.88c). Eye width and eye height measurement methodology is described in 83E.4.2. All counter-propagating signals shall be asynchronous to the co-propagating signals using Pattern 5 (with or without FEC encoding), Pattern 3 or a valid 100GBASE-R signal. Patterns 3 and 5 are defined in Table 86-11. For the case of Pattern 3, with

Draft Amendment to IEEE Std 802.3-2012
IEEE P802.3bm 40 Gb/s and 100 Gb/s Fiber Optic Task Force

IEEE *Draft* P802.3bm/D2.0
14th November 2013

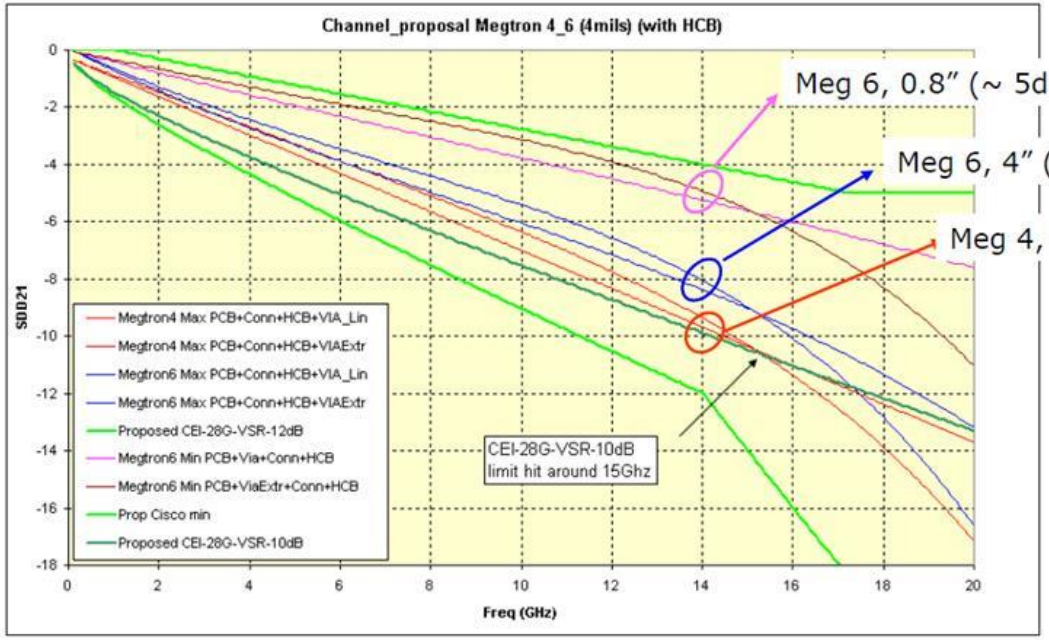
83E.3.1.5 Transition time

The transition times (rise and fall times) are defined in 86A.5.3.3.

83E.3.1.6 Host output jitter and eye height

Host output eye width is greater than 0.46 UI. Host output eye height is greater than 95 mV. Figure 83E–9 depicts an example host output jitter and eye height test configuration. Host output jitter and eye height are measured at TP1a using compliance boards defined in 83E.2. The host output eye is measured using a reference receiver with a continuous time linear equalizer defined in 83E.3.1.6.1. Eye contour measurement methodology is described in 83E.4.2. All counter-propagating signals shall be asynchronous to the co-propagating signals using Pattern 5 (with or without FEC encoding), Pattern 3 or a valid 100GBASE-R signal. Patterns 3 and 5 are defined in Table 86-11. For the case of Pattern 3, with at least 31 UI delay between the

Loss Budget Summary



- 10dB budget is close to meeting the requirements for a 4" host trace
  - ~ 2 dB margin when using Meg 6
  - ~ 0 dB margin when using Meg 4

- Our NEBS/DVT tests assume 2C/min temperature ramp variation, so we should cover these variations with the proposed adaptation method.

- Our first estimation is +/- 1dB variation with respect to the previous CTLE gain setting with a minimum 1Hz frequency rate, but we're open to discussion.

2. Technical Concerns if marginal parts will be adopted (case 1)

Potential link broken at start-up when an IEEE compliant host is mated to a compliant module.

Example – case of a single point of compliance for host (TP1a) and module (TP1).
For some port cases, the host is characterized and qualified to be TP1a compliant with just one CTLE coefficient (e.g. 5dB).
The module is also characterized and qualified to be compliant to a stressed input receiver (TP1), only when CTLE is set to 4dB is given to it.
Host write 5, module expect 4 -> link is broken, should have worked with adaptive mode.

Module's RX compliant CTLE coefficient

Host setting

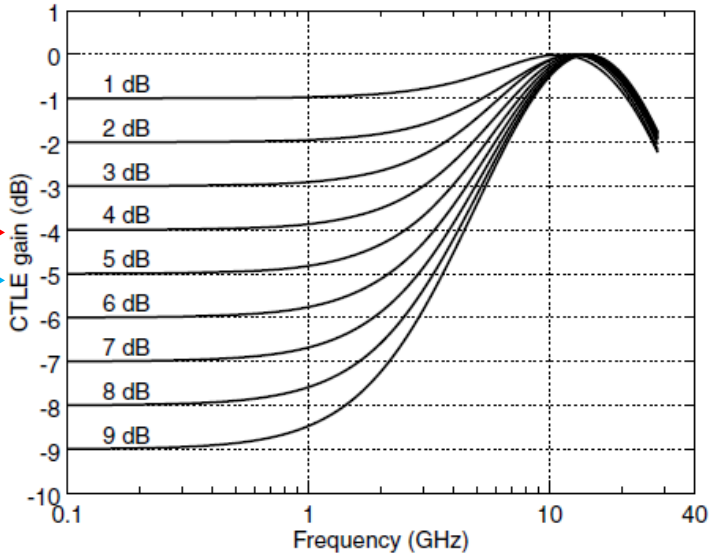


Figure 83E–10—Selectable continuous time linear equalizer (CTLE) characteristic

Neither host and module are responsible for the link broken -> responsibility is shared.

2. Technical Concerns if marginal parts will be adopted (case 2)

Potential link broken at start-up when compliant host is mated to a compliant module.

Example – single point of compliance, case of module and hosts CTLE coefficient mismatch.
For some port cases, the host is characterized and qualified to be TP1a compliant with one CTLE coefficients (e.g. 5 dB).
The module is also characterized and qualified to be compliant to a stressed input receiver, when CTLE is set to 5dB, too.
The way in which the host and module's manufacturer verify compliance makes 1dB mismatch between the two methods - the host should have a 6dB coefficient (instead of 5dB) to the module, to make it working.
(e.g. TP1a host characterization is done leveraging a reference receiver which will typically implement a CTLE filter mathematically , thus an ideal filter. This could lead to discrepancies with a real chip CTLE implementation, as reference receiver will not capture non-idealities and design tradeoffs built into the receiver. Besides this we have to take into account equipment tolerances, etc..),
The host write 5 (optimum value) to the module -> link is broken, should have worked with adaptive mode.

Module's RX compliant CTLE coefficient

Host setting

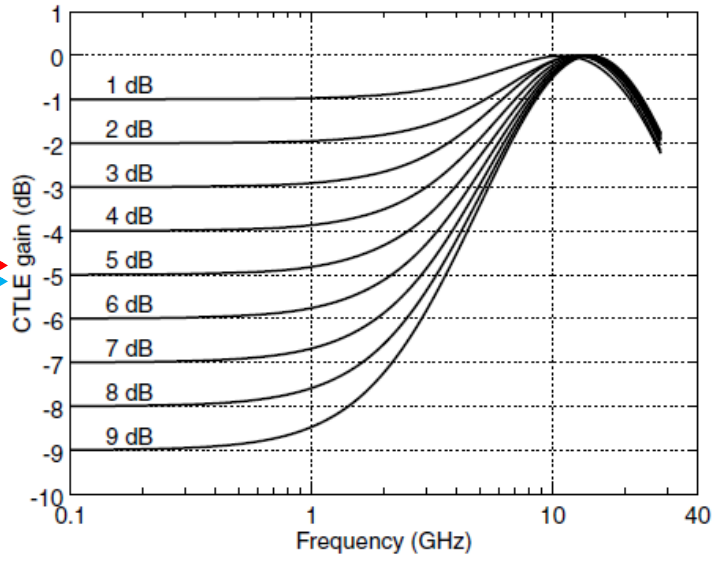


Figure 83E–10—Selectable continuous time linear equalizer (CTLE) characteristic

Responsibility still shared.