White Paper: CMIS-Based Out-of-Band Messaging for Link Training

OIF CMIS-LT White Paper

OIF-CMIS-LT-01.0

May 9, 2024

Abstract:

This white paper provides background on link training and then defines a message catalogue by which any transmit SerDes can be trained initially (or tuned adaptively while in service) based on the needs of the corresponding receive SerDes, with the help of message exchange.

The white paper promotes out-of-band (OOB) messaging via the ubiquituous CMIS management link as a flexible solution for exchanging the link training messages. The technical specification both of the messaging facility and of the representation of the messages will be defined in the projected CMIS-LT supplement. Potential CMIS-LT applications are CEI-112G VSR, XSR, XSR+, and MR, IEEE 802.3 AUI C2M links, Fibre Channel, InfiniBand, OTN, etc.

CMIS-Based Out-of-Band Messaging for Link Training:

Document Editor:

Ali Ghiasi	Marvell

Document Co-Editors:

David Cassan	Alphawave
Michael Klempa	Alphawave
Cathy Liu	Broadcom Inc.
Kent Lusted	Intel
Whay Lee	Marvell
Todd Rope	Marvell
Sridhar Ramesh	MaxLinear

The OIF is an international non profit organization with over 150 member companies, including the world's leading carriers and vendors. Being an industry group uniting representatives of the data and optical worlds, OIF's purpose is to accelerate the deployment of interoperable, cost-effective and robust optical internetworks and their associated technologies. Optical internetworks are data networks composed of routers and data switches interconnected by optical networking elements.

With the goal of promoting worldwide compatibility of optical internetworking products, the OIF actively supports and extends the work of national and international standards bodies. Working relationships or formal liaisons have been established with EA, IEEE 802.3, INCITS T11, Infiniband, IPEC, ITU SG-15, PCI-SIG, SNIA SFF.

For additional information contact:
OIF
5177 Brandin Ct, Fremont, CA 94538
510-492-4040 info@oiforum.com

www.oiforum.com

Contents

List of Tables

List of Figures

Glossary†

ADC: An analog-to-digital converter is a system that converts an analog signal into a digital signal.

AFE: An Analog Front End (AFE) is the analog portion of a circuit which precedes slicer or A/D convertor.

ASIC: An application-specific integrated circuit is an integrated circuit (IC) customized for a particular use, rather than intended for general-purpose use.

AUI: Attachment Unit Interface (AUI) is a physical and logical interface originally defined in the IEEE 802.3 10BASE5 Ethernet and now being used for the C2M applications.

BER: Bit Error Ratio is the number of bit errors divided by the total number of transferred bits during a studied time interval.

C2M: Chip-to-module, a type of AUI defined in IEEE 802.3 similar to the OIF CEI VSR specifications.

CDR: Clock and data recovery, a component that re-establishes the timing of a signal that may have been degraded due to impairments on a transmission line, the retimed signal is now able to continue further to its destination.

CEI: Common Electrical IO, an OIF Implementation Agreement containing clauses defining electrical interface specifications.

CMIS: Common Management Interface Specifications.

CMIS-LT: CMIS-LT denotes (the combination of) a specific message set and a message exchange mechanism based on a CMIS managed link, which may be used for out-of-band link training or tuning applications that require link training entities in receiving and transmitting link endpoints to interact by message exchange. CMIS-LT denotes (the combination of) a specific message set and a message exchange mechanism based on a CMIS a management link, which may be used for out-of-band link training or tuning applications between a pluggable module and host. CMIS-LT provides only the messaging building blocks for out-of-band link training or tuning. Note that any particular link training protocol between these endpoints is not in the scope of CMIS-LT.

CMIS-LT is also the shorthand name for the CMIS supplement specification that defines the CMIS data structures and CMIS interactions (READ, WRITE, Interrupt) implementing the CMIS-LT message set and message exchange. This 2nd meaning can be distinguished by context or by adding the word specification (CMIS-LT specification). The 1st meaning can be distinguished by context or by adding the word messaging (CMIS-LT messaging).

CMOS: A complementary metal-oxide semiconductor (CMOS) is the semiconductor technology used in most of today's integrated circuits (ICs).

CTLE: Continuous time linear equalizer.

DFE: Decision feedback equalizer. A DFE is a nonlinear equalizer that uses past symbol decisions to eliminate the ISI caused by previously detected symbols on the current symbol being detected.

DME: Differential Manchester Encoding.

DSP: Digital signal processor – device using digital signal processing functions to recover data from an impaired signal.

FEC: Forward error correction gives a receiver the ability to correct errors without needing a reverse channel to request retransmission of data.

FFE: Feed forward equalizer.

FIR: Finite impulse response (FIR) is a filter an impulse response settles to zero in finite time.

FLR: Frame loss ratio defined as the ratio between the number of frames that have errors and the total number of frames observed at the MAC/PLS service interface.

HCB: Host Compliance Board, a test board to measure the host output or inject signal into the host input.

HostIn: Host Input (host/ASIC SerDes receiver)

HostOut: Host Output (host/ASIC SerDes transmitter)

IC: Integrated Circuit

ISI: Intersymbol interference.

KR: [A family of] IEEE 802.3 [Physical Layer] specifications using BASE-R encoding over an electrical backplane, based on the CR SerDes. The SerDes capability is similar to the OIF CEI-LR.

LT: Link Training.

LR: Long reach, CEI LR specifies backplane/midplane and copper cable electrical interfaces.

MLSD: Maximum likelihood sequence detector, a detection algorithm for noisy ISI channels.

MCB: Module Compliance Board, a test board to measure the module output or inject signal into the module input.

ModHsIn: Module Host Side Input (module SerDes receiver)

ModHsOut: Module Host Side Output (module SerDes transmitter)

ModMsIn: Module Media Side Input (module optical input TP3)

ModMsOut: Module Media Side Output (module optical output TP2)

MR: Medium reach. CEI MR specifies chip-to-chip electrical interface.

PAM: Pulse amplitude modulation, a form of signal modulation where the message information is encoded in the amplitude of a series of signal pulses. For optical links it refers to intensity modulation.

PAM4: Pulse amplitude modulation-4 is a two-bits modulation that takes two bits at a time and maps the signal amplitude to one of four possible levels.

PCS: IEEE 802.3 Physical Coding Sublayer, sublayer responsible for encoding and FEC.

PMD: IEEE 802.3 Physical Medium Dependent, sublayer consisting of the transceiver or connector for the physical medium.

Preset: A register configuring host (HostOut) or module (ModHsOut) SerDes transmitter to a specific set of FFE taps that maybe defined by IEEE 802.3 or OIF PLL.

Optimum Preset: Host (HostOut) or module (ModHsOut) SerDes transmitter FFE taps on a trained link maybe stored into user Optimum Preset register for future reuse after the link reset.

SerDes: A Serializer/Deserializer is a pair of functional blocks or transceiver commonly used in high speed communications that converts parallel data to serial data and vice versa. The transmitter section is a parallel-to-serial converter, and the receiver section is a serial-to-parallel converter.

SerDes IP Core: A Serializer/Deserializer integrated circuit silicon macro.

SNDR: Signal-to-noise-and-distortion ratio is a measurement of the purity of a signal.

SNR: Signal-to-noise ratio.

SQM: Signal Quality Metric provide a measure of receive signal to noise SNR and BER.

TWI: Two Wire Interface, an example of TWI is I^2C commonly used to manage optical modules.

MCU: Micro-controller, small low power type of CPU that manages optical modules or SerDes IP.

UI: The bit period is commonly called the Unit Interval (UI) when describing an eye diagram.

VEC: Vertical Eye Closure (VEC) is a measure of the ratio of the ideal eye opening (separation between the average levels surrounding the eye) to the measured eye height reported in dB.

VEO: Vertical Eye Opening (VEO) is a measure of eye-opening amplitude at target.

VSR: Very short reach. CEI VSR specifies chip-to-module electrical interface.

XSR: Extra short reach. CEI XSR specifies die-to-optical engine (D2OE) and die-to-die (D2D) electrical interface.

XSR+: Extended Extra short reach. CEI XSR+ specifies die-to-optical engine (D2OE) and die-to-die (D2D) electrical interface.

† Some definitions include content from www.wikipedia.com

1 Introduction and Overview

1.1 Background and Context

The purpose of **link training** (LT), of one direction of a physical link, is to optimize parameters of a signal shaping filter in the upstream transmit SerDes, based on requests from the downstream receive SerDes (of any type: adaptive or not).

Historically, the need for link training came about from receive SerDes implementations utilizing a decision-feedback equalizer (DFE) relying on the upstream transmitter's feed-forward equalizer (FFE) for pre-cursor compensation. Some of today's DSP-based receivers commonly used for copper cable or backplane applications may not require transmit signal shaping, but they may still require lowering signal power. Less capable XSR/XSR+ and VSR receive SerDes will likely need some transmit FFE tuning. Today's VSR/C2M interfaces are part of multi-segment (4-6 segments) links, where each of the segments may require training.

In **in-band** link training solutions the information to be exchanged is transmitted as overhead within the signal used for data transmission. It depends on the specifics of transmission format and protocol, and it requires dedicated hardware.

For instance, the Ethernet in-band link training solution specified in IEEE 802.3 Clause 136/162 [2] is used for point-to-point links between two IEEE 802.3 PCS instances, where continuously exchanged training frames carry the link training information between the two end points. It was developed for point-to-point links across backplanes or copper cable links between systems, and it may not support OIF non-Ethernet links.

In **out-of-band** link training (OOB LT) a separate message communication channel is used for the link training related communication between the two link end points.

For the case of host-module link training, the management interface link is always present, and – as will be shown – it can be used to emulate the required message exchange facility.

A link training **message catalogue** defines the messages available to a physical layer link training solution that uses OOB message exchange between link end points. This offers the potential of being useable for various situations and signals when a generic transmitter and signal model is used.

1.2 Purpose of this White Paper

This **white paper** aims at defining a generic **message catalogue** for transmit FFE based link training applications, at a **conceptual** level, without the details of data representation.

This conceptual message catalog builds on the concept of Ethernet in-band link training, except that the messages are sent out-of-band. It provides enhancements and more flexibility by supporting jumping taps, joint optimization, and requests to store and recall FFE taps.

1.3 CMIS based Host-Module Link Training Support

For an actual link training application, apart form the link training algorithms, approaches, or procedures used by the endpoints, two essential implementation ingredients are then still needed: a message exchange service and a representation (encoding) of the message data.

To enable out of band host-module electrical link training applications while using an ubiquitous CMIS [1] management connection for the exchange of link training messages, the OIF is running a CMIS Host-Module Link Training Support project which is tasked to generate a suitable CMIS supplement called CMIS-LT.

CMIS-LT will specify both the **implementation of messaging** in terms of CMIS management operations (register access and interrupts) and the **detailed message representations**. CMIS-LT will thus enable the implementation out of band host module link training applications, but it will not define the link training procedures employed by the end points.

Potential application of CMIS-LT in link training solutions are OIF CEI 112G-VSR [7], CEI 112G-XSR [7], CEI 112G-XSR+ [8], CEI 112G-MR [7], CEI 112G-Linear [6], and IEEE 802.3 AUI C2M and C2C links [3] for improved SNR and BER.

1.4 White Paper Overview

This white paper is written by SerDes experts and provides some background on SerDes and equalization principles. Understanding equalization principles is not a pre-requisite to understand the concept of CMIS-LT link training. The white paper provides background on the application space, Ethernet in-band link training, host SerDes architecture, equalization, module CDR/DSP architecture, CMIS-LT message catalog for out of band link training, example link training algorithm using CMIS-LT messaging catalog.

2 Link Training for VSR/C2M Applications

2.1 VSR/C2M Application Overview

Error! Reference source not found. shows an example of one potential VSR/C2M implementation block diagram. The basic application consists of the host, the high-speed IO's and channel, and the two-wire interface (TWI) bus. Today, there is no standard defined to implement link training on VSR/C2M links operating at Ethernet 106.25 Gbps or OIF 112 Gbps per lane.

2.2 The Benefit of Link Training

Link training of host-module and module-host paths overcomes measurement inaccuracies and captures the impact of host line cards, manufacturing, and PCB temperature/humidity variations. VSR/C2M links with CMIS-LT are expected to operate with greater margin and would need less debugging.

2.2.1 Benefit of Out of band Link Training

Out of band link training removes the dependency on a specific transmission protocol and the need for dedicated hardware. Link training in essence consist of exchange of messages between the end points on a link. The message exchange requires a communication channel. In the context of host to module and module to host link training, one may use the CMIS management communication link to transport the link training messages. The message catalog and facility to transmit these messages via CMIS called CMIS-LT.

2.3 VSR/C2M Transmit FFE Tuning

Figure 1 shows a Host Compliance Board (HCB) that has been inserted into a host system to measure TP1a output or TP4a stress input. With an HCB plugged into the host, the HostOut FFE taps are adjusted while Vertical Eye Closure (VEC) and Vertical Eye Opening (VEO) are measured at TP1a for VEC below its limit while eye opening VEO remain above its limit. Host system HostOut FFEs are tuned for optimum setting with an HCB on a scope and during operation optical module are plugged into the host and expected to operate with the same host static FFE setting.

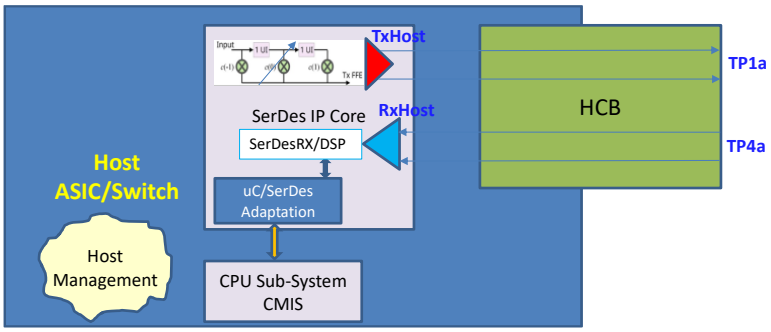

Figure 1: Host Measured with HCB

Figure 2 shows a module that has been inserted into a Module Compliance Board (MCB) to measure TP4 output or TP1 stress input. While plugged into the MCB, its ModHsOut FFE taps are adjusted while VEC and VEO are measured at TP4 with addition of a short synthetic channel and long synthetic channel for maximum eye opening VEO and minimum penalty VEC. In the IEEE Std. 802.3-ck, the ModHsOut FFE has two module output modes (AUI-S and AUI-L) corresponding to the relevant host electrical interface channel.

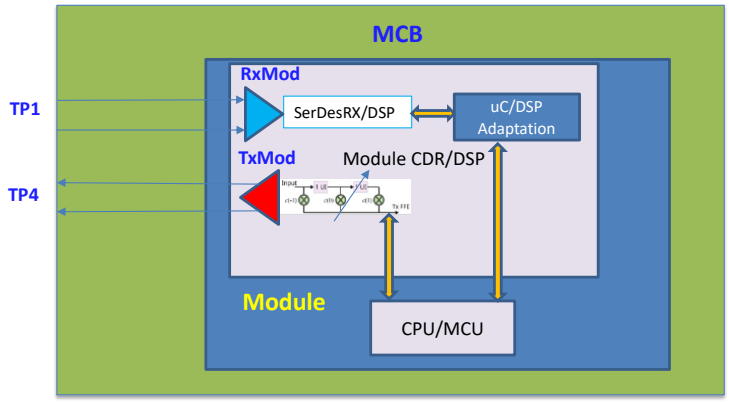

Figure 2: Module Measured with MCB

2.4 Limitation of VSR/C2M Fixed Transmit FFE Settings

Current IEEE Std. 802.3ck-2022 [3] and OIF 112G-VSR [7] some implementers may configure the host transmitter at TP1a, as shown in Figure 1, tuned during manufacturing to a fix transmitter FFE setting per system port and assumed these fixed transmit FFE settings are optimum for any optical module over the life of the product. Optical modules transmitter at TP4a, as shown in Figure 2, tuned during manufacturing to two possible settings AUI-S (short) or AUI-L (long), where module is expected these two transmit FFE settings are sufficient to cover all possible host channels.

There is no standard way to implement link training; unless systems designers implement their own proprietary methods, the performance of the SerDes can be negatively impacted by environmental changes, process variation and product life degradation. To meet the BER target for 112G/lane with a single fixed HostOut or ModHsOut FFEs the system may need to find other methods to ensure adequate performance, and this may include proprietary training algorithms or the allocation of excess margin.

Having a single transmitter setting at TP1a and two settings (AUI-S and AUI-L) at TP4a has margin and yield impacts. Another major point of discrepancy is the HCB or MCB behavior versus real system connectors and channels. The difference between the signal measured

through HCB or MCB terminated into 50 Ω scope versus a module plugged into a host system can be very significant.

2.5 Relationship of CMIS-LT with IEEE in-band LT

CMIS-LT messaging catalog allow to provides equivalent function to IEEE in-band link training [2] techniques. CMIS-LT message catalog allow to optionally train (i.e., 112G-VSR/C2M) and tune the host's (i.e., HostOut) electrical interface TxFFE/amplitude and/or the module's (i.e., ModHsOut) electrical interface TxFFE/amplitude without requiring integration of link training hardware into the SerDes/DSP ASIC.

CMIS-LT messaging protocol like IEEE 802.3 CL136/162 also include presets registers, but with additional capability to store TX FFE optimum settings into a preset register for future use. Additional CMIS-LT messaging capability is the ability to tune transmitter amplitude based on receive SerDes feedback. With prior optimum TX FFE stored, the CMIS-LT capable host may apply the optimum preset value as starting point for LT optimization and in some cases LT optimization may not be necessary as the prior TX FFE settings remain at optimum.

3 Equalization for 112G

In a high-speed serial link system, transmitter (e.g., HostOut, ModHsOut) and receiver (e.g., HostIn, ModHsIn) equalizations normally work together to compensate the channel loss and cancel the inter symbol interference (ISI) [9]. The block diagrams of a serial link system including a transmitter, a channel and a receiver are shown in Figure 3 and Figure 4, with an analog based receiver and a digital signal processing (DSP) based receiver, respectively.

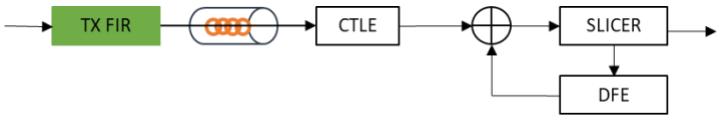

Figure 3: Transmitter and receiver equalization block diagram for an analog based receiver

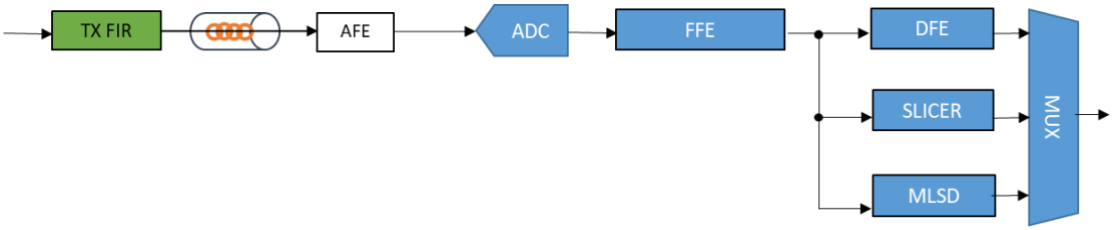

Figure 4: Transmitter and receiver equalization block diagram for a DSP based receiver

To achieve a good trade-off between performance and power/cost, transmitter equalization and receiver equalization need to work together. The pulse responses of a CEI-112G VSR channel in Figure 5 illustrates how the channel ISIs are cancelled by transmitter (TX) FIR, receiver CTLE, and DFE sequentially. In general, TX FFE is implemented as a multiple-tap FIR filter to mitigate pre-cursor and post-cursor ISIs.

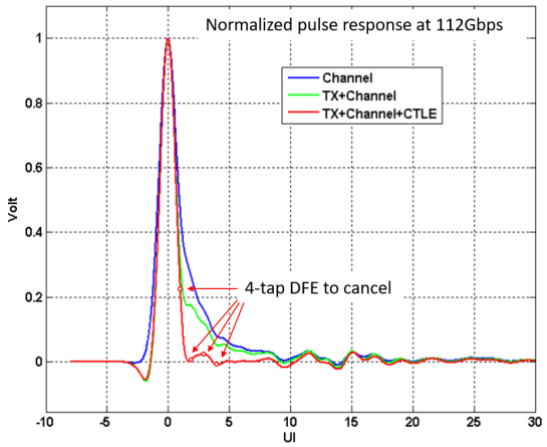

Figure 5: Normalized pulse responses of a CEI-112G-VSR channel

The co-optimization between transmitter and receiver equalizations becomes indispensable with the growth of link rate and channel loss [2]. Such co-optimization is normally achieved through link training between the host and module devices in case of KR/CR in IEEE Std. 802.3ck-2022 [3]. However, the current OIF CEI-112G-VSR/XSR(+) [7]/[8], IEEE Std. 802.3ck-2022 [3], and IEEE P802.df [4] have not defined link training specifications for chip-to-module and die-to-OE interfaces. Figure 6 shows an example of the performance differences between optimal transmitter FFE settings, which can be achieved with transmitter optimization, and 802.3ck module output modes (AUI-L and AUI-S addressing long and short channels) over variable C2M channels; the penalty from the optimal transmitter FIR settings could be as large as 1.6 dB VEC and 40% VEO.

https://www.ieee802.org/3/ck/public/21_07/ghiasi_3ck_01_0721.pdf

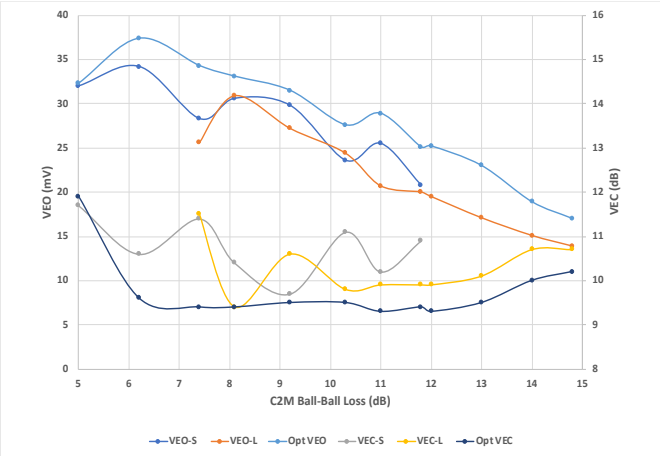

Figure 6: Performance differences between optimal Tx FFE settings and preset settings

3.1 Range of Transmitter FFE

Table 1 lists transmitter FFE tap range and steps for CEI-112G-VSR, CEI-112G-XSR and CEI-112G-XSR+ interfaces [7]/[8]. CEI-224G and IEEE 802.3dj [5] projects are still being defined but one can expect more transmitter FFE pre-cursor taps, and larger ranges are required, which results in higher demand for transmitter FFE optimization or/and link training.

Table 1: Transmitter FFE specifications over CEI-112G VSR and XSR (+) interfaces

Interface	Pre/Post taps	Tap step size	Tap range and step size [1]
CEI-112G-VSR [1]	2/1	0.02	C(-2)=0:0.1, C(-1)=-0.2:0, C(0)=0.65:1.0, C(1)=-0.1:0
CEI-112G-XSR	1/1	0.02	C(-1)=-0.14:0, C(0)=0.6:1.0, C(1)=-0.18:0
CEI-112G-XSR+	2/3	0.02	C(-2)=-0.04:0.04, C(-1)=-0.18:0, C(0)=0.55:1.0, C(1)=-0.18:0, C(2)=-0.1:0.04, C(3)=-0.06:0.04

1. C(0) is main tap, C(-) are the pre-cursor taps, and C(+) are the post-cursor taps.
2. In case of VSR the FFE taps are the recommendation for the channel.

4 Host SerDes IP Core Overview

CMOS scaling has enabled tremendous information processing on a single host ASIC die. But ultimately, data needs to move between different chips, and this is the role of SerDes IP. To transmit and receive high-speed data at minimum power and area overhead, a typical host ASIC, for example a 25.6Tb/s high-speed Ethernet switch, will include 256 lanes of 100 Gbps/lane SerDes.

To meet the demands of the latest high-speed networking technologies, such as OIF-CEI 112G [7], IEEE Std. 802.3ck-2022 [3] and especially looking forward towards 200 Gbps/lane electrical signaling, SerDes design has evolved to flexible and robust digital DSP based architectures. Channel responses may include significant energy far away from the main cursor caused by reflections. As the symbol rate has increases, the unit interval (UI) gets reduced, and cursor energy due to reflection in unit intervals moves away from the main cursor. This increased Inter-Symbol Interference (ISI), combined with PAM4 signaling, and other impairments such as crosstalk and noise, degrade the signal integrity to the point where more complex DSP based equalization and forward-error-correction (FEC) are necessary to meet acceptable SQM performance.

4.1 Detailed Block Diagram and Operation of SerDes IP Core

Figure 7 shows an example SerDes IP Core block diagram and the anatomy of a modern SerDes.

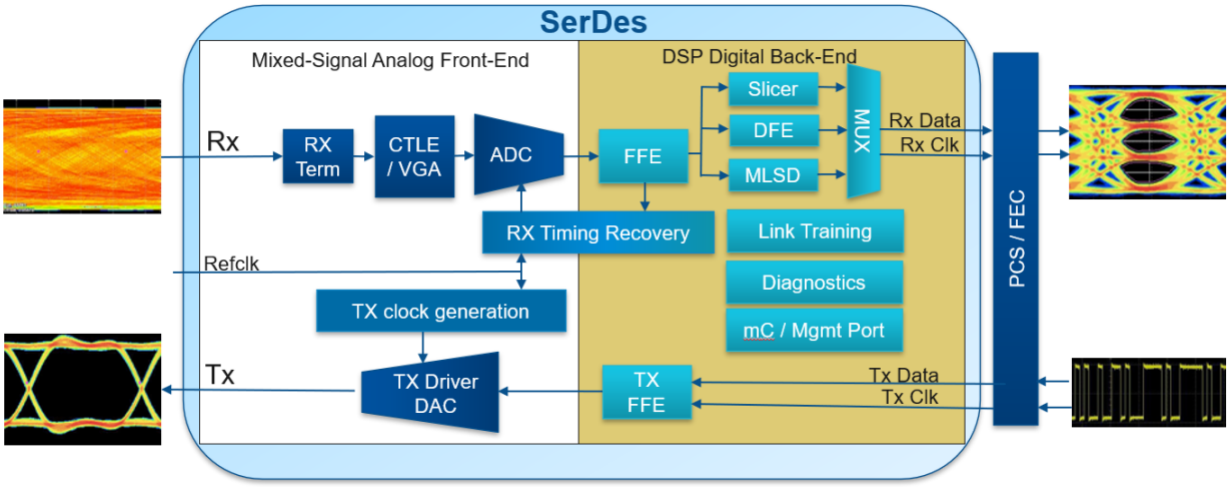

Figure 7: An Example SerDes IP Core Block Diagram

A SerDes comprises some, or all, of the following building blocks:

- RX On-Chip Termination – provides broadband differential termination to match the characteristic impedance of the wireline channel and ESD protection.
- Continuous-Time Linear Equalizer (CTLE) – Provides high-frequency 'boost' to equalize a portion of the channel response before the ADC and DSP equalization.
- Variable Gain Amplifier (VGA) – Provides wide bandwidth programmable gain to ensure that the input signal approaches the full-scale range of the ADC for best performance.

- RX Analog to Digital Converter (ADC) – converts an analog signal into a digital code to allow for further Digital Signal Processing (DSP) before the slicing decision into a PAM4 data stream.
- RX Timing Recovery – A Baud rate Clock and Data Recovery (CDR) loop comprising a timing error detector, loop filtering, and RX clock phase/frequency adjustment to find the optimal ADC sampling point of the signal.
- RX Feed Forward Equalizer (FFE) – a symbol spaced finite impulse response filter that contains multiple pre-cursor and post-cursor taps to equalize the channel response.
- Decision Feedback Equalizer (DFE) – Takes the samples received and subtracts off the estimated ISI from the previous symbol (the 1+αD partial response).
- Maximum Likelihood Sequence Detector (MLSD) - enables extra-long channels by using residual ISI to make an informative guess at the bit sequence using an LMS algorithm.
- TX Clock Generation – Common implementations include a low-phase noise PLL to enable low-noise clock generation and independent rates on all Tx and Rx lanes.
- TX Feed Forward Equalizer (FFE) – A symbol spaced finite-Impulse response filter to provide de-emphasis on the transmitted waveform.
- TX Driver DAC – A transmitter implementation that utilizes digital to analog converter to convert the digitally equalized data stream with variable amplitude analog signal output with back termination that drives the channel.
- Link Training (LT) – Link training is a mechanism through which the receiver can request to adjust the partner's transmitter to optimize performance.
- Diagnostics – test features including data loop backs, PRBS pattern generators and checkers, and Signal Quality Monitor (SQM) monitors allow SerDes self-testing and channel monitoring.
- µC / Management Port – An internal micro-controller and management control port to the host ASIC micro-controller to access all SerDes features and interact with the SerDes link training logic.

Although outside the SerDes, Forward Error Correction (FEC) is employed to allow correct interpretation of signal with relatively poor SNR to prevent even more complex equalization from being necessary but still meet the system frame loss ratio (FLR).

All of these blocks have correlated parameters that need tuning and can all positively or negatively affect other blocks and the overall SQM of the system. In an ideal scenario, a system's behavior can be characterized by its frequency response from the transmitter through the channel to the receiver. Where the combined SerDes TX and RX equalization would implement the inverse of the combined channel frequency response. In practical implementations, the channel is unknown, and due to temperature and other affects, can be changing over time. The noise and crosstalk environment also affects equalization; it is often desirable to equalize to a partial response of the target channel frequency response, to avoid undue noise amplification in linear equalizers. This is especially true for SerDes receivers that include MLSD. Modern SerDes employ various adaptive equalization strategies to deal with this reality. The current Link Training (LT) approach adopted in IEEE Std. 802.3-2022 [2] and IEEE 802.3ck [3] tunes the link when it is brought into service.

5 Transceiver Module Overview

An example of a pluggable transceiver based on QSFP-DD module form factor is shown in Figure 8 (figure is courtesy of QSFP-DD MSA). The pluggable transceiver can be connected into a high-speed switch (or other networking device), providing high-speed connectivity. The transceiver module assembly is comprised of:

- Module Micro-Controller: Controls and monitors module components – CDR/DSP IC, TOSA, ROSA, and PMIC, querying various registers in a periodic fashion, or via interrupts. It also maintains a link to the host (via TWI) for module management (via CMIS [1]).
- CDR/DSP IC: Performs the major PAM4 signal processing; equalization, detection, signal conditioning/pre-emphasis. It also provides access to signal quality metrics for diagnostics for link monitoring & maintenance.
- TOSA: Generates the light signal and performs electrical to optical translation in transmit path. Includes laser, and modulator.
- ROSA: Performs optical to electrical translation in receive path. Includes Photodetector, and Transimpedance Amplifier.
- Power management unit: Controls power supplies (including bias voltages, through bias networks) to active devices in the module.

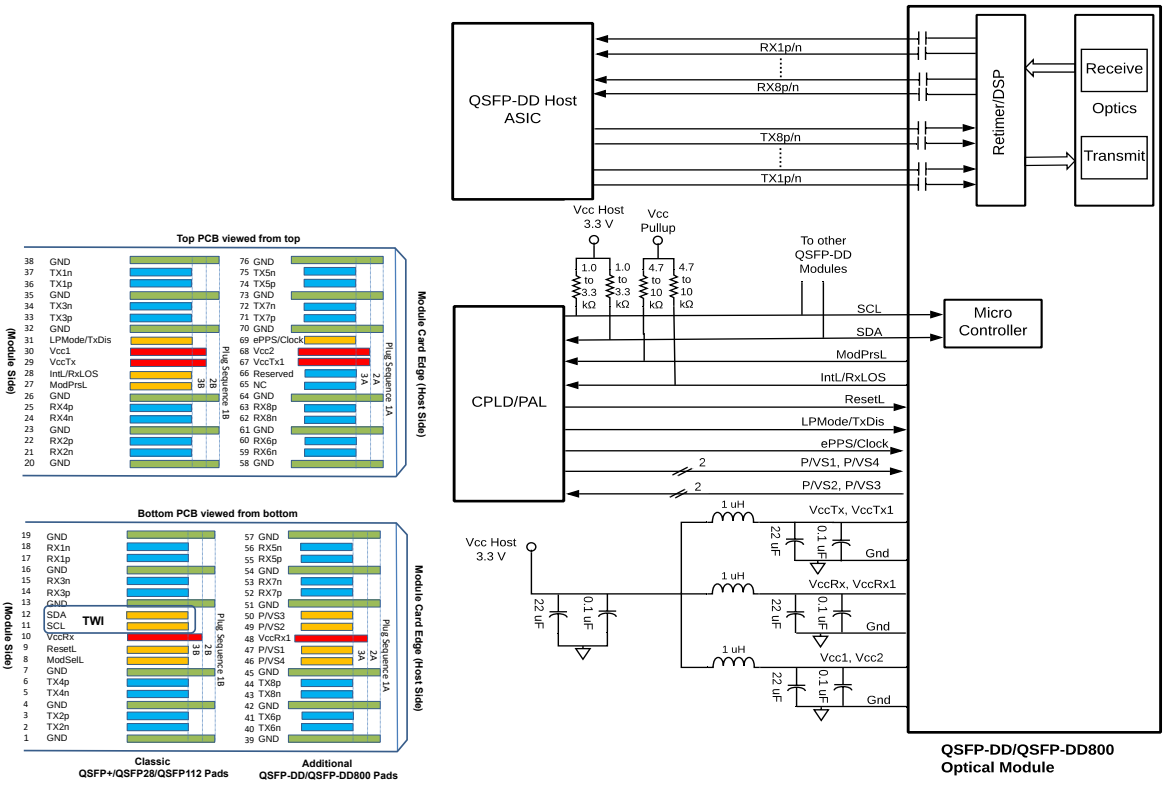

Figure 8: QSFP-DD Pluggable Transceiver Module Interface and Example Block Diagram

5.1 Detailed Block Diagram and Operation of CDR/DSP IC

Example CDR/DSP IC block diagram shown is shown in Figure 9 is for IMDD (Intensity Modulation Direct Detection), CMIS-LT is also applicable for other signaling such as coherent. Notable features are:

- Media and Host-Side Receivers provide some or all of:
  - Adaptive analog/digital equalizers to correct for channel loss/ISI
  - Histogram and SNR monitors on each RX lane
  - Non-linear slicers to translate incoming signal to PAM symbol bits
- Media and Host-Side Transmitters provide some or all of:
  - Analog/digital pre-equalizers to pre-condition TX signal
  - Non-linear mapping of PAM symbol bits to signal levels
- Control/mgmt. interface to communication with module micro-controller
  - Interrupts for real-time communication & control
  - Software/firmware over TWI interface to module micro-controller for monitoring/diagnostics.

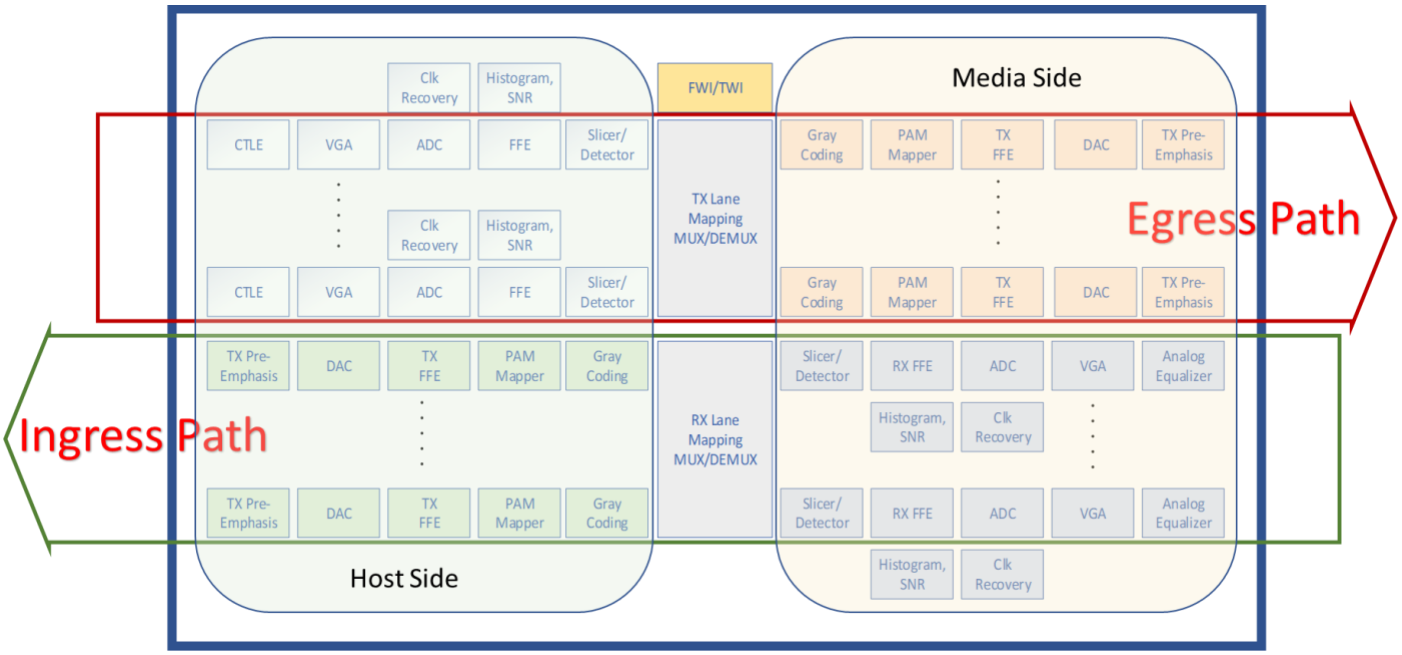

Figure 9: Functional Diagram of CDR/DSP IC

It is seen that the general functions of the host-side and media-side signal processing blocks are very similar. However, a key distinction should also be observed. The TWI (electrical) interface is dedicated for management purposes. The management interface can be used to implement an out of band signaling channels and CMIS-LT. No such out-of-band communication channel exists on the media side interfaces to the remote link endpoint.

5.2 Modern Optical Transceiver Module Software Stack

A typical module management software stack running on a micro-controller embedded within a module is shown in Figure 10.

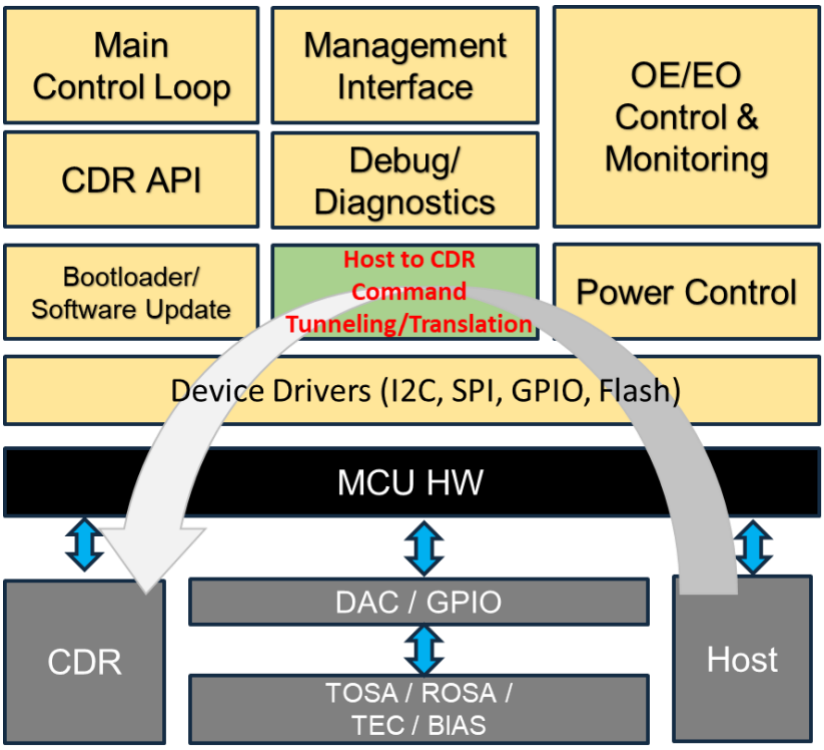

Figure 10: Module Management Software stack

Salient features are outlined below.

- Module micro-controller runs control software for all embedded functions
  - OE/EO bias, gain, TEC
  - CDR/DSP control (e.g., loopback) and monitoring
- CDR/DSP IC provides diagnostic and debug information to module micro-controller
  - SNR, slicer histograms, etc.
  - TX and RX Equalization state (CTLE, FFE)
  - PAM slicer levels, thresholds
- Module micro-controller may implement “tunnel” from switch to CDR/DSP IC
  - Tunneling allows switch ASIC to access more and newer CDR/DSP features and diagnostics without requiring module SW upgrades.

6 Ethernet Link Training for High Speed SERDES

The IEEE Std. 802.3 Standard for Ethernet defines several link training protocols for backplane and copper cable interconnects based on various SERDES signaling rates. These link training protocols are part of the Ethernet physical layer PMD (Physical Medium Dependent) sublayer's control function. The Ethernet link training protocol uses an in-band messaging procedure in which a receiver can communicate status as well as make requests to the other end to tune for optimal performance, such as to change transmitter equalizer settings. The link training protocol in the IEEE Std. 802.3 [2] can be found in Clause 72.6.10 (for 10 Gbps/lane rates), in Clause 92.6.10 (for 25 Gbps/lane rate), in Clause 136.8.11 (for 50 Gbps/lane rates) and in Clause 162.8.11 (for 100 Gbps/lane rates).

6.1 Principles of Ethernet Link Training

In Ethernet, the link training protocol is performed in-band using the data channel between the end points. The link training protocol uses a specific training frame structure and process to exchange information between the two endpoints. The training frame structure consists of a frame marker followed by a controlled channel and a training pattern. The Ethernet training frame structure used for 50 Gbps/lane and 100 Gbps/lanes rates are shown in Figure 11.

The frame marker is composed of a unique identifier to denote the start of the training frame. The control channel consists of two fields: a control field and a status field. The control field provides requests to the partner for action while the status field reports responses and current status of the local receiver. The training pattern comprises a PRBS pattern to represent the mission mode data patterns. The frame marker and the control channel are constructed of cells, where each cell is eight-unit intervals in length, to improve frame decoding before the receiver is fully adapted.

Figure 11: Training frame format

For Ethernet link training, this unique frame structure is distinct from mission mode data and is sent continuously during the training process. The exchange of the training frames continues as both sides adapt to the channel's characteristics, evaluate in the appropriate receiver metrics, and adjust their partners transmitter configuration as required.

At a high level, the link training process consists of the following steps:

I. Perform the necessary SERDES initialization and resets
II. Start sending training frames
III. Detect training frames from the partner
IV. Link training operation

a. Adapt the local receiver using the training frames from the partner
b. Act on requests from the partner received in the incoming control field
c. Make requests to the partner to adjust its transmitter

V. Repeat step 4 until the local receiver meets the required performance metrics (eye height, signal-to-noise ratio, channel pulse response, etc.)
VI. Wait for partner to complete its receiver adaptation, as needed
VII. Transition to mission data.

The local receiver can make two types of coefficients change requests to the transmit SerDes via the control field portion of the link training frame. The first request type is an initial condition request, also known as a preset. The preset is a specific set of predefined TxFFE values and ratios that enable to receiver to jump quickly to better settings. Typically, there are several presets defined to represent various channel losses. The second request type is individual coefficient control, where the receiver asks for a specific coefficient to move in a specific manner. The individual control method enables a receiver to make fine tune adjustments to maximize performance metrics.

The partner responds to requests via the status field portion of the link training frame. Several valid response groups are currently defined in the IEEE Std. 802.3-2022, including:

– Not updated – the requested change has not been made or no changes were requested
– Updated – the requested change was made
– At limit – the requested change reached a limit of allowed values or ratios.

6.2 Detailed Operation of Ethernet Link Training

To better understand the concept of Ethernet link training, it is instructive to focus on Steps 4 and 5 by looking at a specific example. In this example, both Host A and Host B are in Step 4 of the link training process and Host B will exchange an individual coefficient control type message (move coefficient C0) to Host A for action.

In Figure 12, the SerDes IP core in Host B (ModHsIn) determined that it needed assistance from Host A (HostOut) to move Coefficient 0 (C0) improve the requirement metric. The SerDes IP core in Host B notifies the local Link Training (LT) engine.

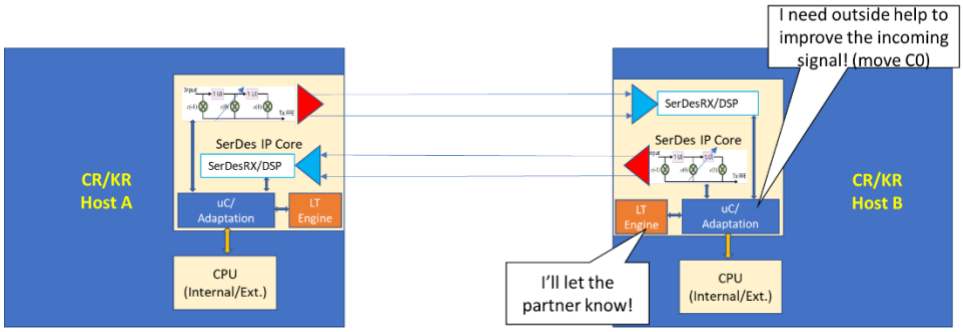

Figure 12: Help needed

In Figure 13, the LT engine in Host B prepares a message by encoding the appropriate information into the control field of the outgoing link training frame.

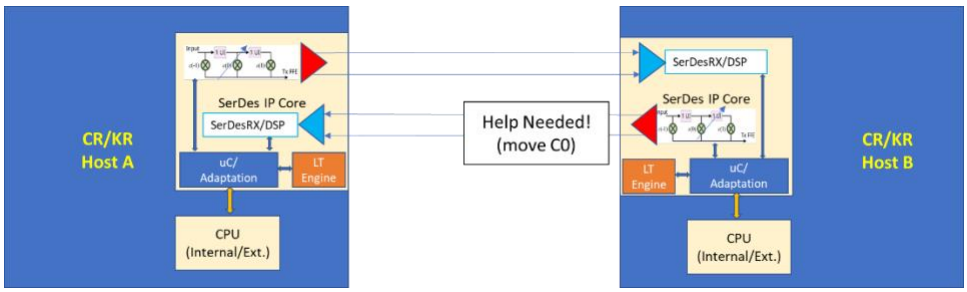

Figure 13: Outgoing message

In Figure 14, the LT engine of Host A receives the message in the link training frame and decodes it.

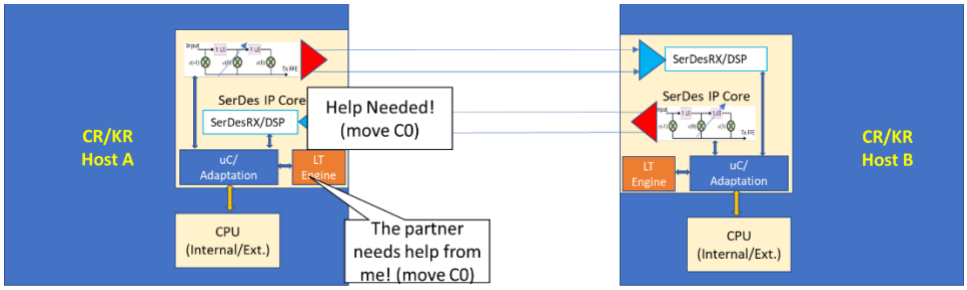

Figure 14: Message received

In Figure 15, Host A's LT engine passes the message to the local SerDes IP Core for action.

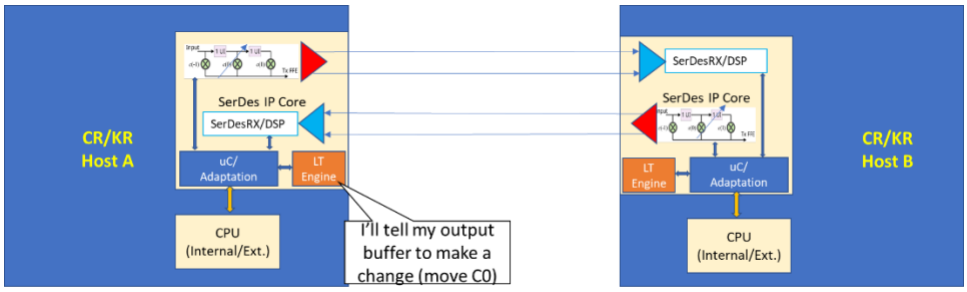

Figure 15: Passing the message

In Figure 16, the transmitter coefficient C(0) of Host A (HostOut) is adjusted as instructed per the message.

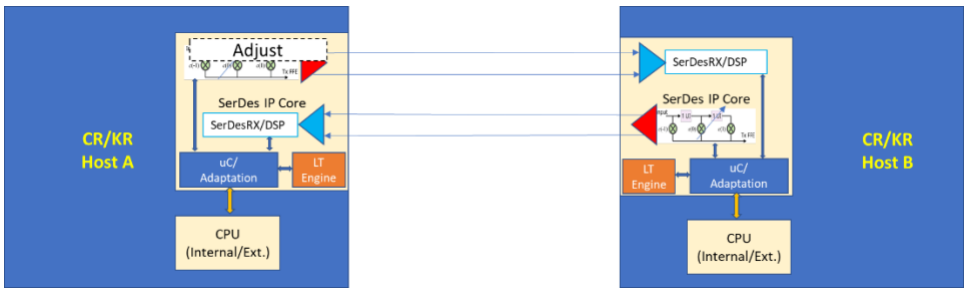

Figure 16: Adjustment made

In Figure 17, the LT engine in Host A prepares a message by encoding the appropriate information into the status field of the outgoing link training frame.

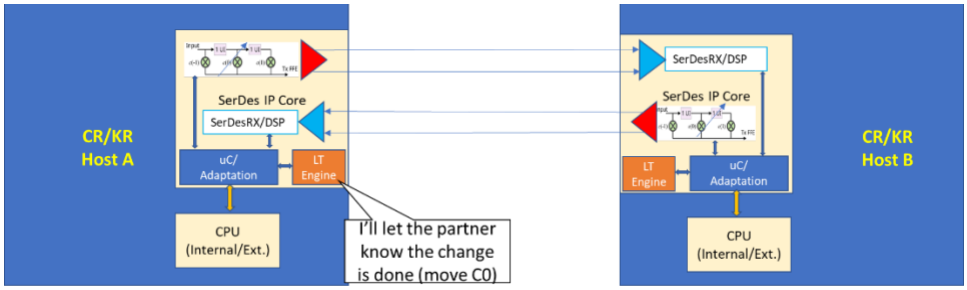

Figure 17: Preparing return message

In Figure 18, the LT engine in Host A acknowledges that the request was executed and reports back the results.

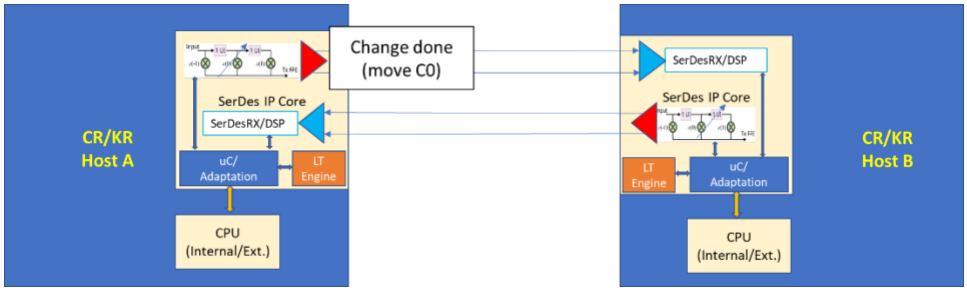

Figure 18: Return message

In Figure 19, the message is received by Host B and passed back to the SerDes IP Core as a completed request.

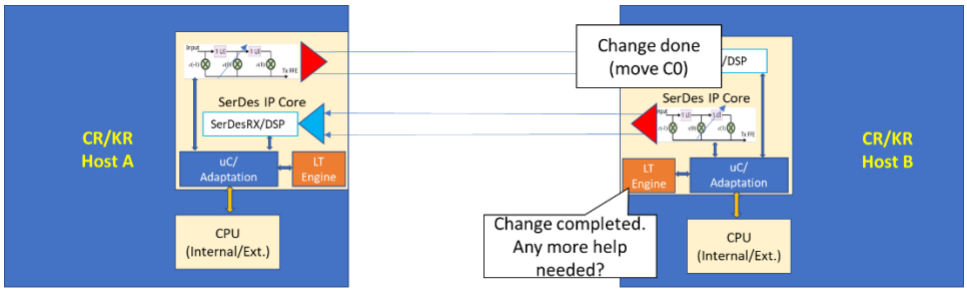

Figure 19: Completed request

If the local SerDes IP core in Host B determines that no further requests from the partner are necessary, it would move to Step 5. Otherwise, it will repeat the procedure illustrated above in Figure 12 through Figure 19 as required to achieve the required as required until the necessary performance metric is achieved as determined by the receive SerDes.

The same type of request for help can occur in parallel in the other direction (from Host A to Host B).

The control field and status field of Ethernet link training as well as the request-response mechanism can easily be adapted for use over an out-of-band interface, such as CMIS-LT.

7 Link Training Architectures with CMIS-LT

CMIS-LT specifies both an out-of-band message catalogue for link training purposes and the messaging mechanism that allow link training entities (engines) in host and module to tune the host's transmitter (HostOut) based on the feedback from the module's receiver (ModHsIn), and, likewise, to tune the module's transmitter (ModHsOut) based on the feedback from the host's receiver (HostIn), see Figure 20.

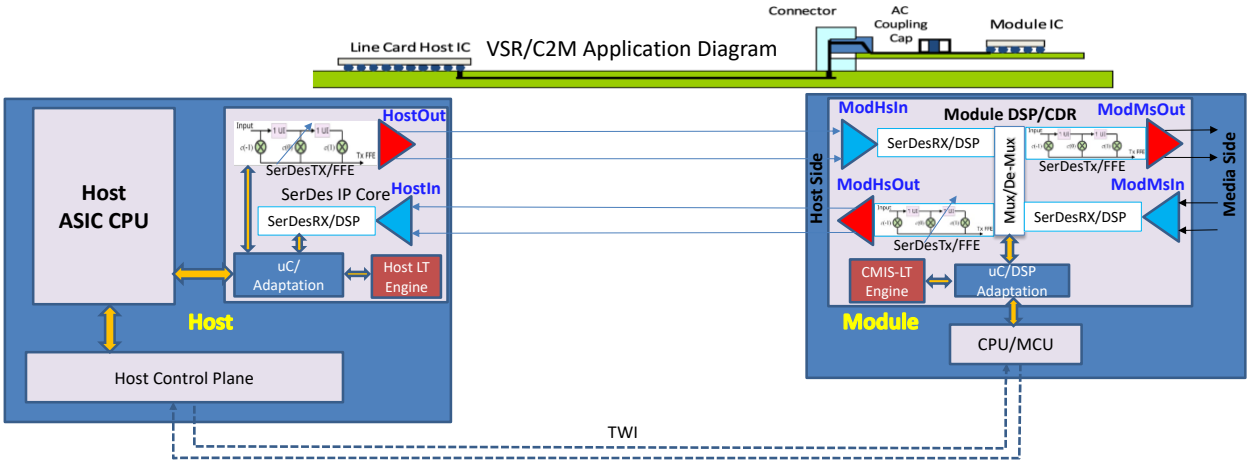

Figure 20: Typical CMIS-LT Application Block Diagram[1]

7.1 CMIS-LT Use Cases

The CMIS-LT messaging facilities over CMIS enable link training solutions for VSR/C2M applications without requiring dedicated hardware. Example potential uses cases for CMIS-LT but are not limited to following list:

- Use CMIS-LT during manufacturing to create generic presets
- Use CMIS-LT to automate manufacturing test
- Use CMIS-LT to create optimum preset for a specific module at deployment
- Use CMIS-LT to train VSR/C2M link
- Use CMIS-LT to read SNR/BER
- Use CMIS-LT to monitor VSR/C2M link degradations.

7.2 Adaption Engines and CMIS-LT Messaging

CMIS-LT based link training implementations will employ existing adaption engines in the host's SerDes IP core and in the module's CDR/DSP, see Figure 20.

[1] CMIS for 1st time will be defining host and mode TX and RX blocks, editors are soliciting input on the TX and RX block names.

The adaptation engine for the host SerDes IP core may typically reside in the SerDes IP core. The CMIS-LT host SerDes adaptation engine communicates to the module through the host control plane via TWI. Requests from the module CDR/DSP LT Engine are send to the module CPU/MCU, the host will then read and act on the requests.

Host software communicates with the SerDes adaptation engine and Management Communication Interface (MCI). Host software tunnels the CMIS-LT requests to/from the SerDes adaptation engine and MCI. .

Adaption engines and the LT engines are self-contained in the host SerDes IP core and Module CDR/DSP, and only the request from the LT engines are relayed to the partner TxFFE through CMIS [1]. CMIS-LT protocol preserves the classic SerDes IP core and Module CDR/DSP partition and responsibilities.

7.3 In-service Tx FFE Adjustment

Some of the less capable and lower power receivers such as XSR/XSR+ receivers may benefit from in-service FFE adaptation in HostOut and ModHsOut, mitigating SerDes performance degradation, which may result from environmental changes.

Messages for in-service Tx FFE adjustments is an optional feature provided in the CMIS-LT message catalogue. In service TxFFE adjustment must be done carefully and with small steps to mitigate potential impact to the data traffic .

8 CMIS-LT Messaging Facility and Message Catalogue

Link training, i.e. the tuning of a transmit FFE by a downstream SerDes receiver in a multi-lane interface is specific to each Egress and Ingress lane.

8.1 CMIS-LT Messaging Facility

CMIS-LT provides out-of-band bidirectional and symmetrical message communication between module (ModHsIn) and host (ModHsOut), and between host (HostIn) and module (ModHsOut).

Messaging from module to host requires some “emulation” because the module is a passive target of management operations in CMIS. The CMIS-LT implementation of messaging from module to host will therefore be based on a CMIS Interrupt that directs the host to read the message data from the module via CMIS READ operations.

Messaging from host to module can be implemented via CMIS WRITE operation sequences as specified in CMIS-LT.

CMIS-LT will also specify which messages are acknowledged (receipt) or replied (execution).

8.2 CMIS-LT Message Catalog

The CMIS-LT message catalog has seven main functional parts:

I. TxFFE presets, see 8.2.1
II. TxFFE tuning, see 8.2.2
III. Amplitude adjustment, see 8.2.3
IV. Pre-coding control, 8.2.4
V. Signal integrity (SQM), 8.2.5
VI. SQM alarm, 8.2.6
VII. Message statuses, 8.2.7.

An illustrative example implementation of how a link training scheme may use messages form the message catalogue is provided in appendix 12.

8.2.1 TxFFE Preset

CMIS-LT commands that control selection of following presets:

1. Optimum presets per lane (TxFFE settings for a given host-module link after initial LT)
2. TxFFE settings defined by IEEE 802.3 or OIF PLL
3. User defined presets.

The host may optionally store the optimized host and module TxFFE’s settings to allow brining up links after reset, possibly avoiding full iterative tuning.

8.2.2 TxFFE

The catalog of CMIS-LT messages controlling HostOut and ModHsOut FFEs will contain messages or message interactions to support the following functions:

4. Advertise TxFFE pre/post taps per lane
5. Increment/decrement TxFFE step sizes
6. Read TxFFE to specific tap values
7. Set TxFFE to specific tap values
8. TxFFE Max/Min Limit indication.

8.2.3 Tx Driver Amplitude

The catalog of CMIS-LT messages to read and adjust host HostOut or module ModHsOut electrical driver amplitudes will contain messages or message interactions to support the following functions:

1. Increment or decrement HostOut/ModHsOut amplitude per lane by one step
2. Read HostOut/ModHsOut current amplitude
3. Request specific HostOut/ModHsOut amplitude
4. HostOut/ModHsOute Max or Min limit reached interrupt.

8.2.4 Tx Pre-Coder

The catalog of CMIS-LT messages to enable or disable pre-coder will contain messages or message interactions to support the following functions:

1. HostOut or ModHsOut precoder on/off per lane.
2. Read HostOut or ModHsOut precoder on/off per lane status.

8.2.5 Signaling Quality

The CMIS-LT message catalog of reading and setting signal integrity target is used by the advance CMIS-LT example implementation, see 13. The use of SQM is for advance user that 1st establish baseline SQM for a given link and DSP, knowing the baseline SQM setting target SQM for the adaptation engine. CMIS-LT signal integrity messages or message interactions to support the following functions:

1. Read current SQM
2. Set SQM target.

8.2.6 SQM Alarm

The catalog of CMIS-LT messages to advertise and set SQM alaram to indicate degradation of the link during operation. SQM alarm supports the following functions:

1. SQM alarm advertisement per lane
2. SQM alarm threshold.

8.2.7 Message Status Catalog

The catalog of CMIS-LT response messages to status back after execution of a task. Message status catalog supports the following functions:

1. FFE update status
2. Set FFE status
3. Set preset status
4. AMP update status
5. Precoder update status
6. Set SQM target status
7. Set SQM status.

9 Integration of CMIS-LT-based Link Training with CMIS

9.1 States

If a CMIS host and a CMIS managed module both support CMIS-LT, the question arises when and in which Data Path states the link training interactions can occur.

To be compliant with CMIS state machine specifications, the host may enable link training in suitable Data Path states. In the parlance of previous sections this amounts to activating the LT engines.

The actual link training interactions governed by the LT engines will have to occur per electrical lane, while the relevant Data Path is in DPInitialized state. In other words, host side link training occurs while the module Tx outputs (media side) are still muted.

9.2 Behavior

While still out-of-service in the DPInitialized state, the host can enable the test patterns required for any particular link training algorithm (defined e.g., by the OIF PLL or IEEE 802.3) and have the LT engines performing the relevant link training interactions.

9.3 The Simplified Example Link Training Implementation

With PRBS generators and checkers enabled in host and module, the host LT engine and the module LT engine cycle through the presets and if any of the presets are good enough the host configures the module for mission mode.

If none of the presets are good enough the host and module LT engines start iterative transmitter adjustment, after several iterative tuning if signal quality is good enough LT exit with success otherwise exit with timeout/failure. For detail operation of simplified host and module LT transaction, see 11.

9.4 The Advance Example Link Training Implementation (optional)

With PRBS generators and checkers enabled in host and module, the host LT engine and the module LT engine cycle through the presets and if any of the presets meets the target SQM the host configures the module for mission mode.

If none of the presets meet the target the host and module LT engines start iterative transmitter adjustment, after several iterative tuning if target SQM is met then LT exit with success otherwise exit with timeout/failure. For detail operation of advance host and module LT transaction, see 12.

10 Summary

CMIS-LT emulates a bi-directional out-of-band message communication channel between transmit SerDes and receive SerDes via a CMIS management interface an defines a message catalogue which allows the partners to tune the link for optimal performance based on SerDesRx/DSP SI evaluation.

Host SerDes IP Core and Module CDR/DSP blocks may have correlated parameters that need adjustment and can all positively or negatively affect other blocks and the resulting SI.
The ideal equalization will be the exact inverse of the channel response. In real implementations, the channel is unknown, the behavior model not clearly defined, and due to varying environmental condition, process, and manufacturing variations link tuning is necessary to ensure the target SerDes SI performance is met over the entire product population.

The OIF PLL Management group is defining an optional messaging protocol called CMIS-LT that allows bi-directional out-of-band communication between the host SerDes and the module CDR/DSP to optimally train host HostOut FFE and the module ModHsOut FFE.
CMIS-LT based out-of-band messaging can support any data transmission format or rate.
Link training provides increased receiver margin for (i.e., 112G) VSR/C2M links. Link training for 224G VSR/C2M is expected to be necessary and may get standardized by OIF PLL and IEEE Std. 802.3dj. By defining a method for the module receiver to communicate with the host transmitter, link training can be performed as efficient and succinct as possible (or determine it's not needed at all) to allow for faster bring up times and have a system that operates with higher margin. CMIS-LT messaging can be optionally engaged after Data path initialization.
CMIS-LT messaging supports preset registers that can be used by system implementers to minimize link training time, based on previously tuned links.

11 Appendix A – A simplified Example LT Implementation

This appendix provides a simplified example of host-module LT and module-host LT implementation. The simplified link utilizes the message catalog [x] and follows the principle of IEEE CL136/162 link training. One enhancement CMIS-LT provide over the IEEE in-band link training is the direct amplitude tuning. The actual host SerDesRX DSP (RxHost DSP) and module SerDesRx DSP (RxMod DSP) may choose different adaptation path from the example shown in this appendix.

11.1 Module to Host path LT

Detailed operation of the module to host LT key steps are listed below and shown in the flowchart Figure 25:

I. Host enables host PRBS Gen
II. Host invokes link training RxMod LT
   - RxMod DSP requests presets from host
     - RxMod DSP cycles through presets
   - Best preset selected
   - RxMod DSP evaluates best preset and if the settings are good enough then may skip iterative tuning
   - RxMod DSP may initiate iterative tuning of HostOut FFE and AMP
     - Module requests HostOut FFE or Amp adjustment
     - Host adjusts HostOut FFE or Amp then informs module update completed
     - If there is room for improvement RxMod DSP will retry iterative tuning if n>1 (n=number of retry) and if good enough then exit tuning
     - If n=0 then exit with LT failure otherwise LT exit with LT success
III. Host turns off PRBS generator and engages mission mode.

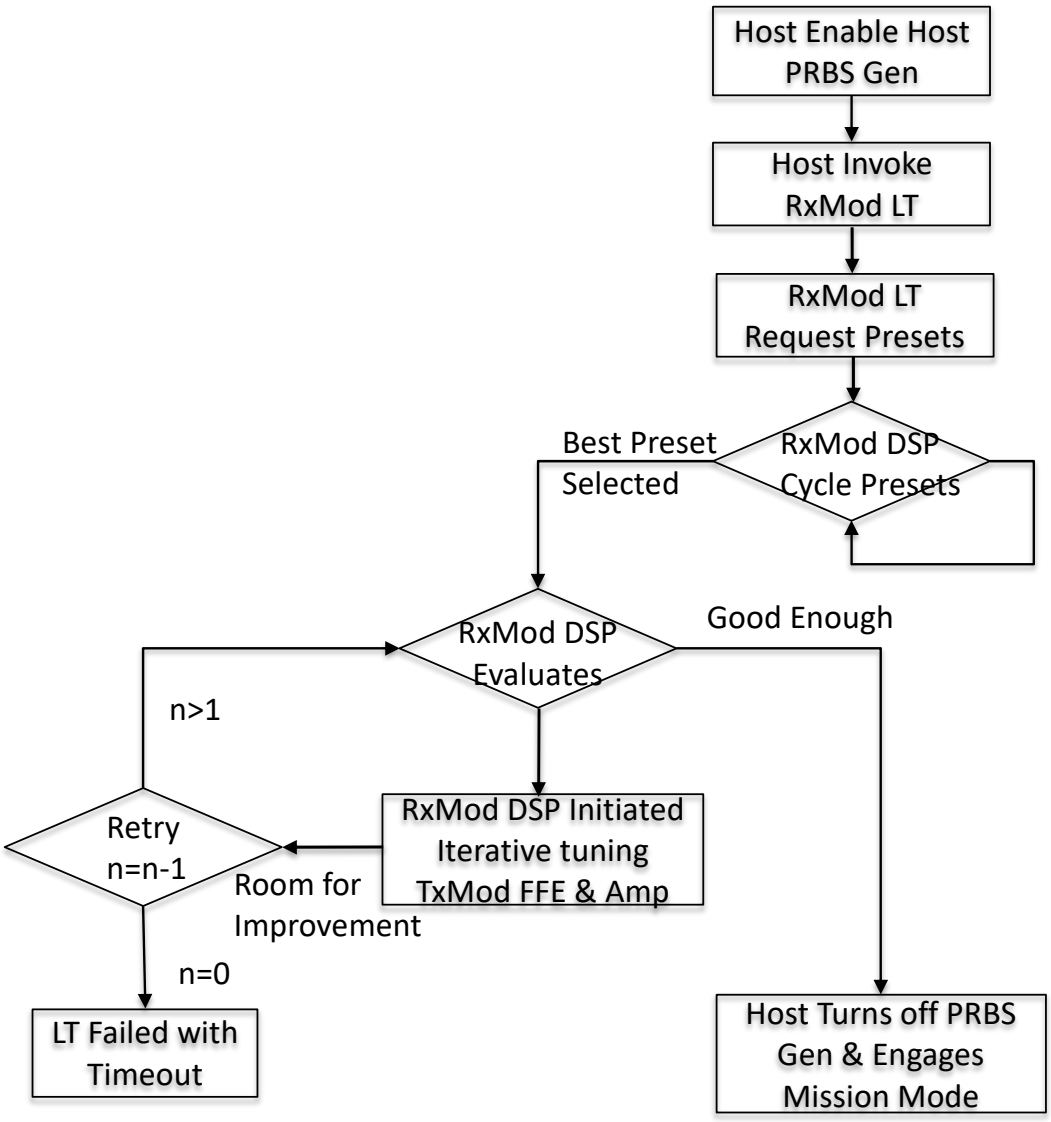

Figure 21: Flow Chart of the Module to Host LT

Host SerDes IP Core LT transactions flow is shown Figure 26. If link training is not completed within the specified number of tries, the receive SerDes will declare link failure due to timeout.

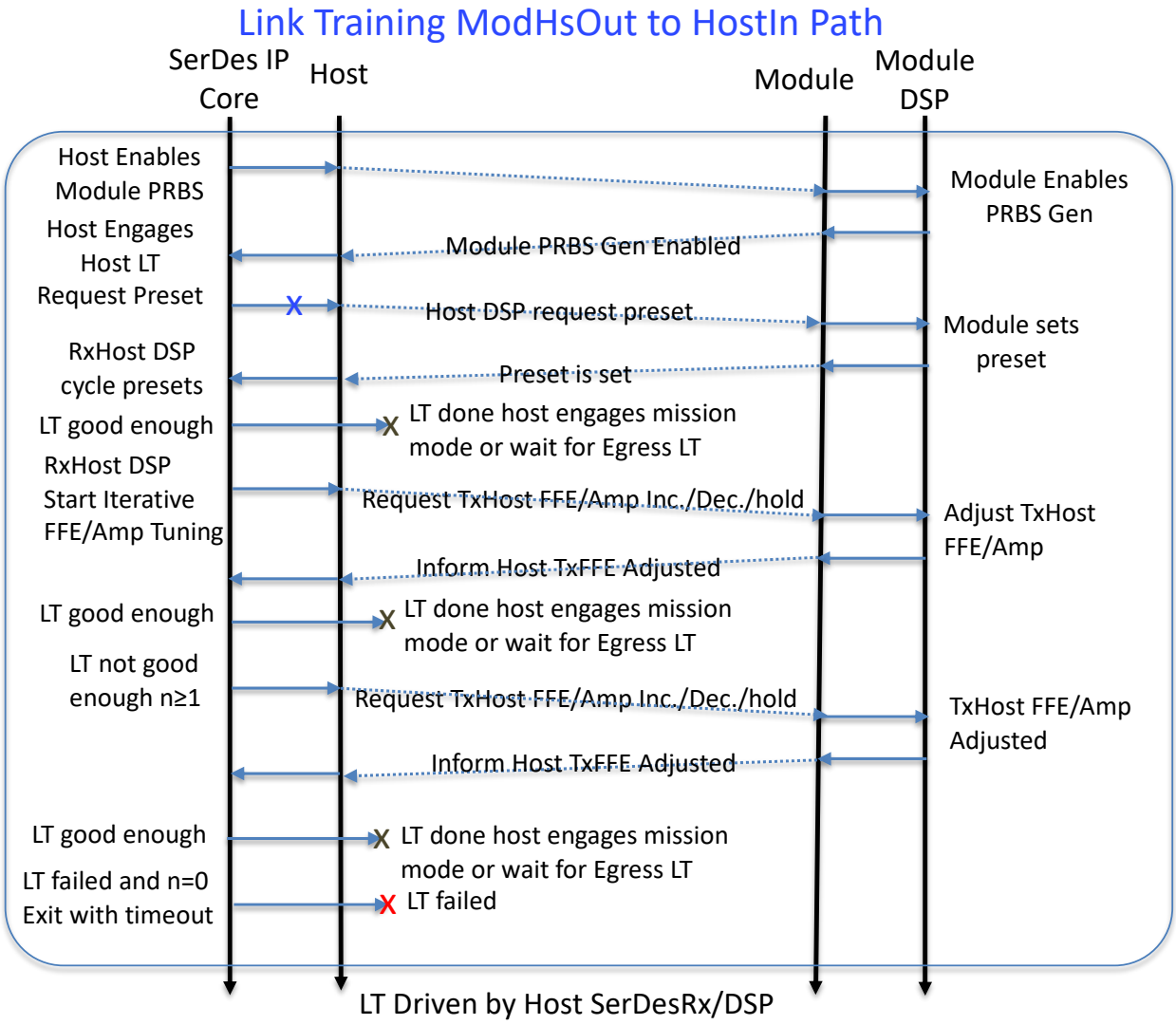

Figure 22: Host SerDes IP Core Initiated LT Transactions

11.2 Host to Module path LT

Operation of the host to module LT is very similar to the module to host LT, except all requests from the module starts with an Interrupt and then host read and tend to the module requests. Interrupt is also used to inform the host if module ModRx DSP is trained or if training failed due to timeout.

Key steps for the host to module LT are listed below and shown in the flowchart of Figure 27:

I. Host enables module PRBS Gen
II. Host invokes link training RxHost LT

- RxHost DSP invokes presets
  - RxHost DSP cycles through presets
- Best preset selected
- RxHost DSP may initiate iterative tuning of ModHsOut FFE and AMP
  - Host requests ModHsOut FFE or AMP adjust
  - Module informs host ModHsOut FFE or AMP updated
  - If there is room for improvement RxHost DSP will retry iterative tuning if n>1 (n=number of retry) and if good enough then exit tuning
  - If n=0 then exit with LT failure otherwise LT exit with LT success

III. Host turns off PRBS generator and engages mission mode.

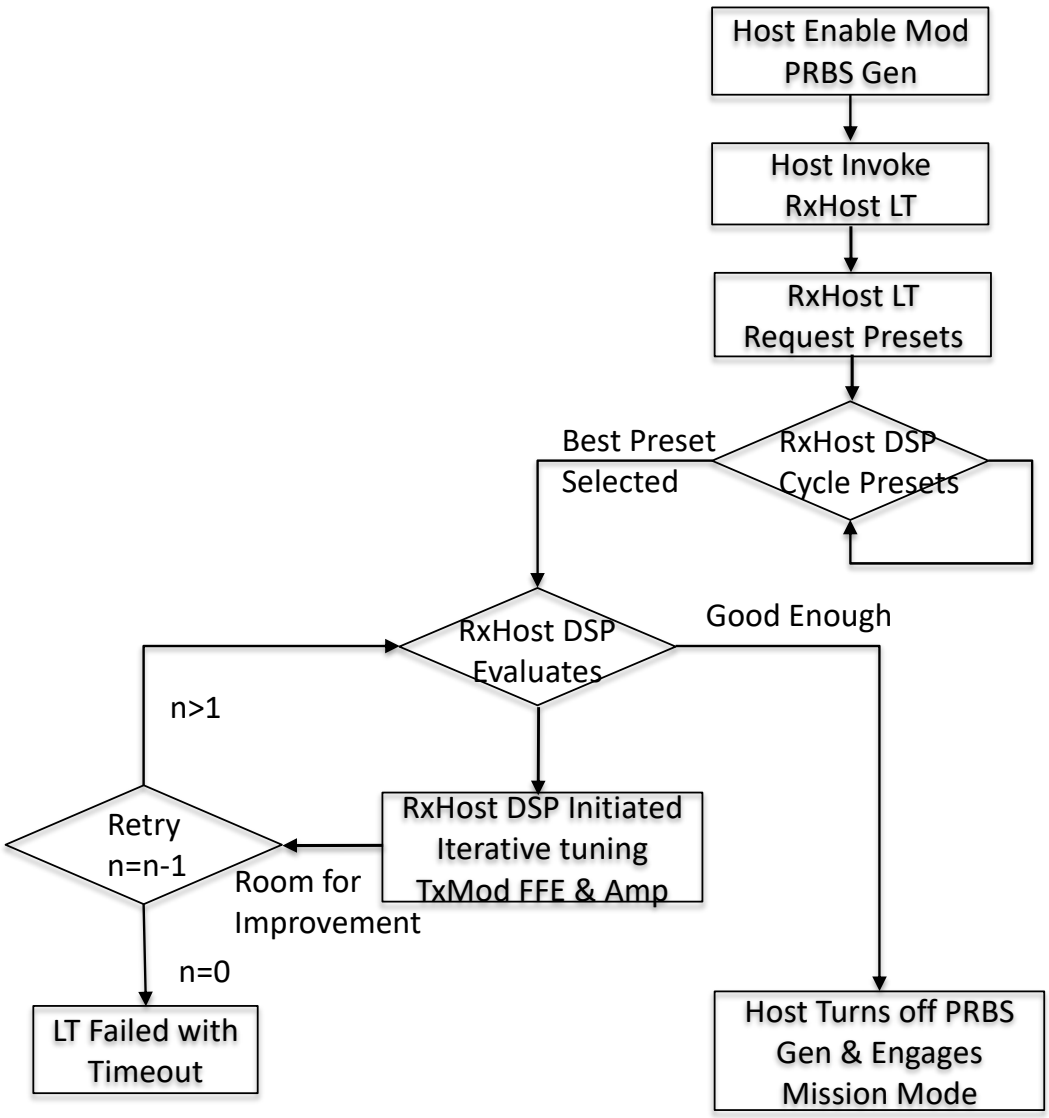

Figure 23: Flowchart of Module to Host LT

Module CDR/DSP initiated LT transactions are shown in Figure 28. If link training is not completed within the specified number of tries, the receive SerDes will declare link failure due to timeout.

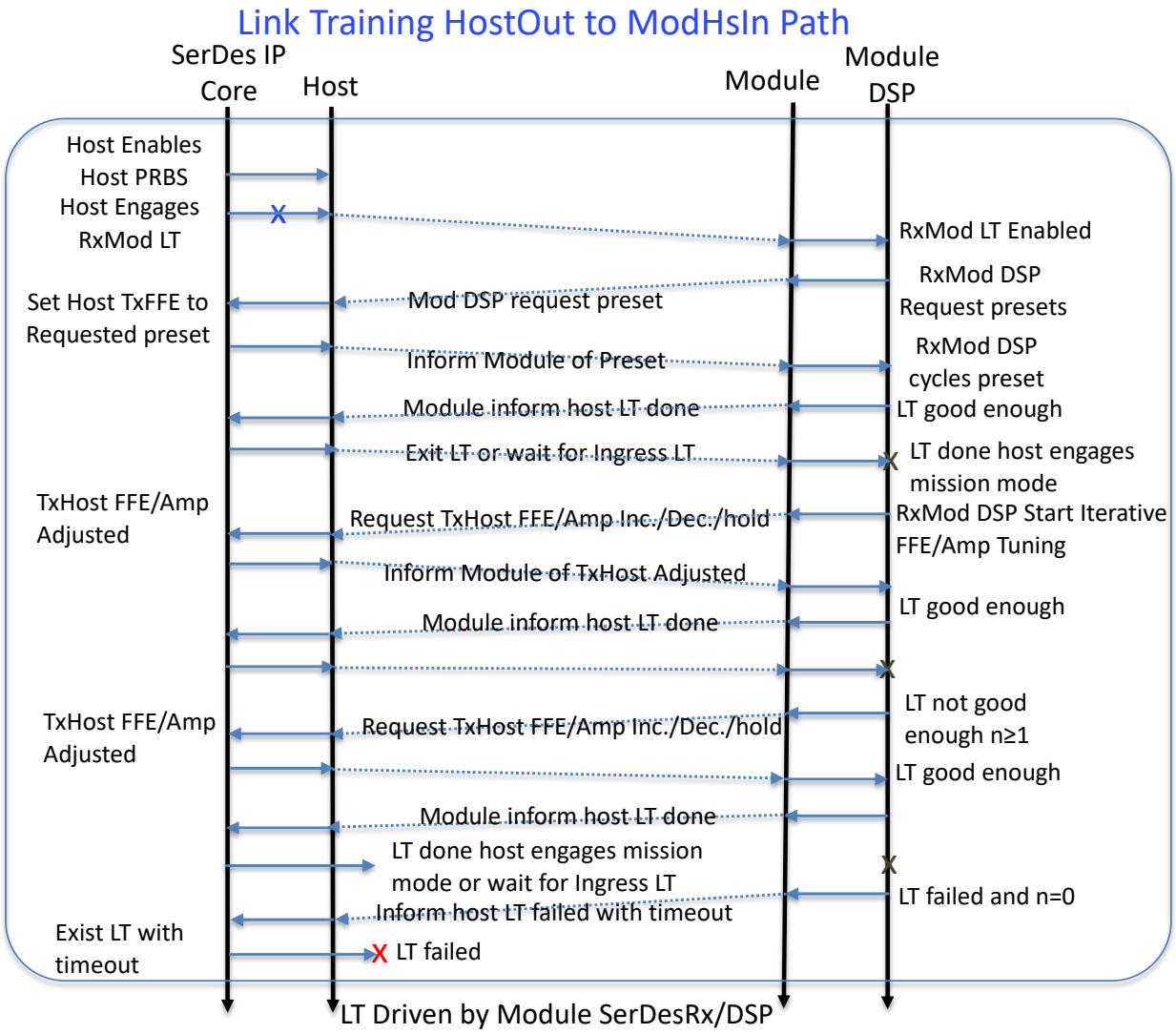

Figure 24: Module CDR/DSP Driven LT Transactions

12 Appendix B – Example LT Advanced Implementation

This appendix provides an advanced example of host-module LT and module-host LT implementation. These examples are provided to illustrate how CMIS-LT messages can be used. The actual host SerDesRx DSP (RxHost DSP) and module SerDesRx DSP (RxMod DSP) may choose different adaptation path from the example shown in this appendix.

12.1 Module to host path LT operation

Detailed operation of the host LT Engine key steps are listed below and shown in the flowchart Figure 25:

I. Based on SI parameters SQM host determines if ModHsOut-HostIn need training with default setting.
   - If SQM target is met, then go to mission mode
   - If tuning is required, host receiver requests the module transmitter to enables ModHsOut PRBS generator as defined in CMIS

II. Host may set target
   - Meet best SQM, meet certain SQM

III. Host starts RxHost DSP link training
   - CMIS host fetches the request from RxHost DSP
   - RxHost DSP may try using presets including enabling pre-coder before iterative tuning or HostIn LT directly goes to iterative tuning and adjustment
   - RxHost DSP SerDes may request from the module its normalized ModHsOut FFE coefficients
   - RxHost DSP SerDes goes through the preset settings
     - Module informs host ModHsOut FFE updated
     - if for any preset setting the target SQM is met then RxHost DSP informs the host
     - Host may store optimized ModHsOut FFE
     - Host may engage the mission mode
   - If the target SQM are not met HostIn SerDes may try iterative tuning (increment/decrement), request brand new set of taps (jump), including increment/decrement amplitude
     - Module informs Host ModHsOut FFE and/or Amplitude updated
     - RxHost DSP continues till receiver determines it is tuned to optimum setting or meet certain SQM and stops and informs the host with success or failure
     - Assuming success, then host engages mission mode and host may store optimized ModHsOut FFE
     - Host may retry optimizing, n=number of retry
     - Target SQM is failed – may have faulty host or module, or the target SQM set too high.

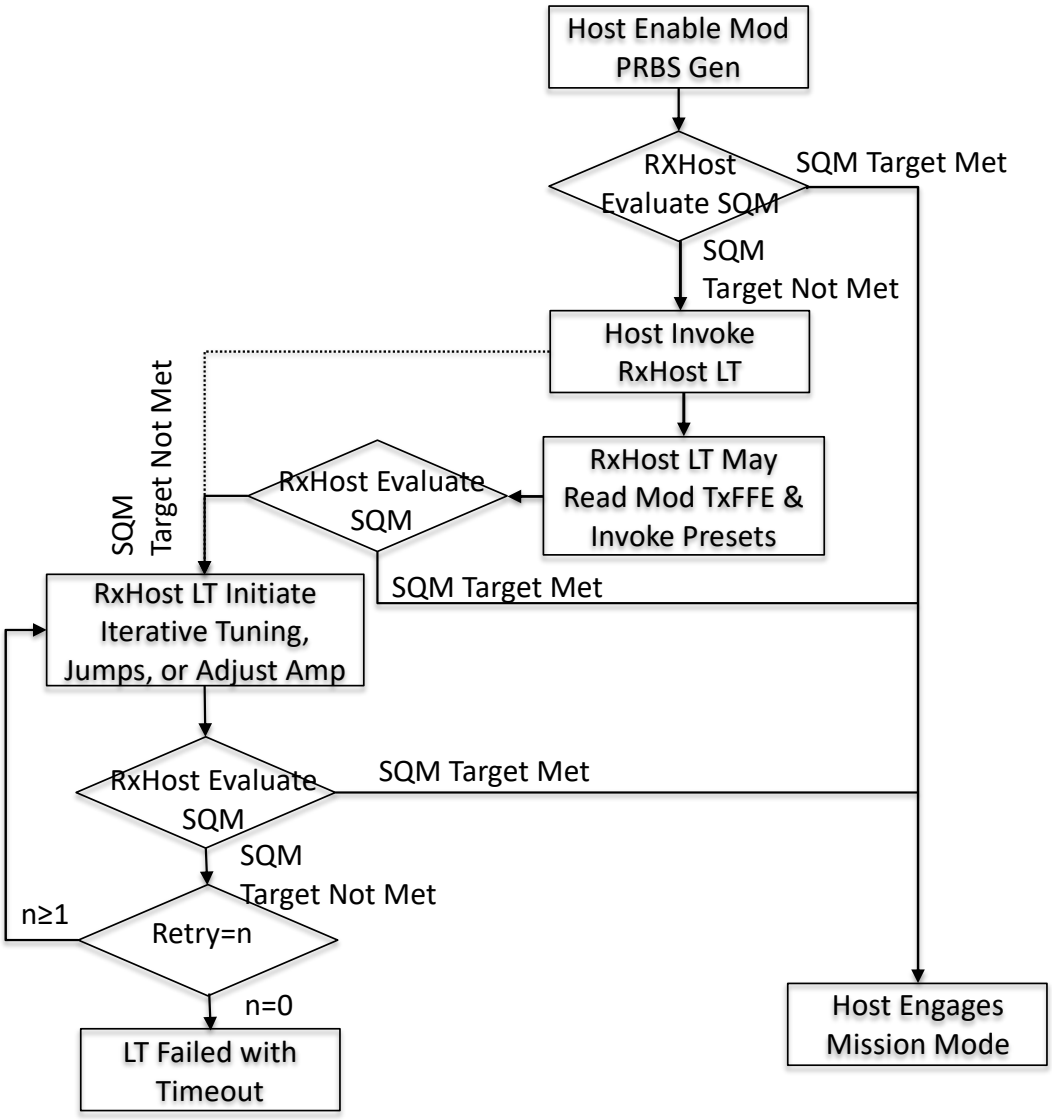

Figure 25: Flow Chart of the Module to Host LT

Host SerDes IP Core initiated LT transactions are shown Figure 26. The transaction flow diagram includes Pre-LT optional SQM evaluation to determine if link training is needed. If SerDes receiver performance metric SQM is met without link training the host may just transition to mission mode indicated by 1st instance of green "X". If link training is not completed within the specified link training time, the receive SerDes will declare a link host failure due to timeout.

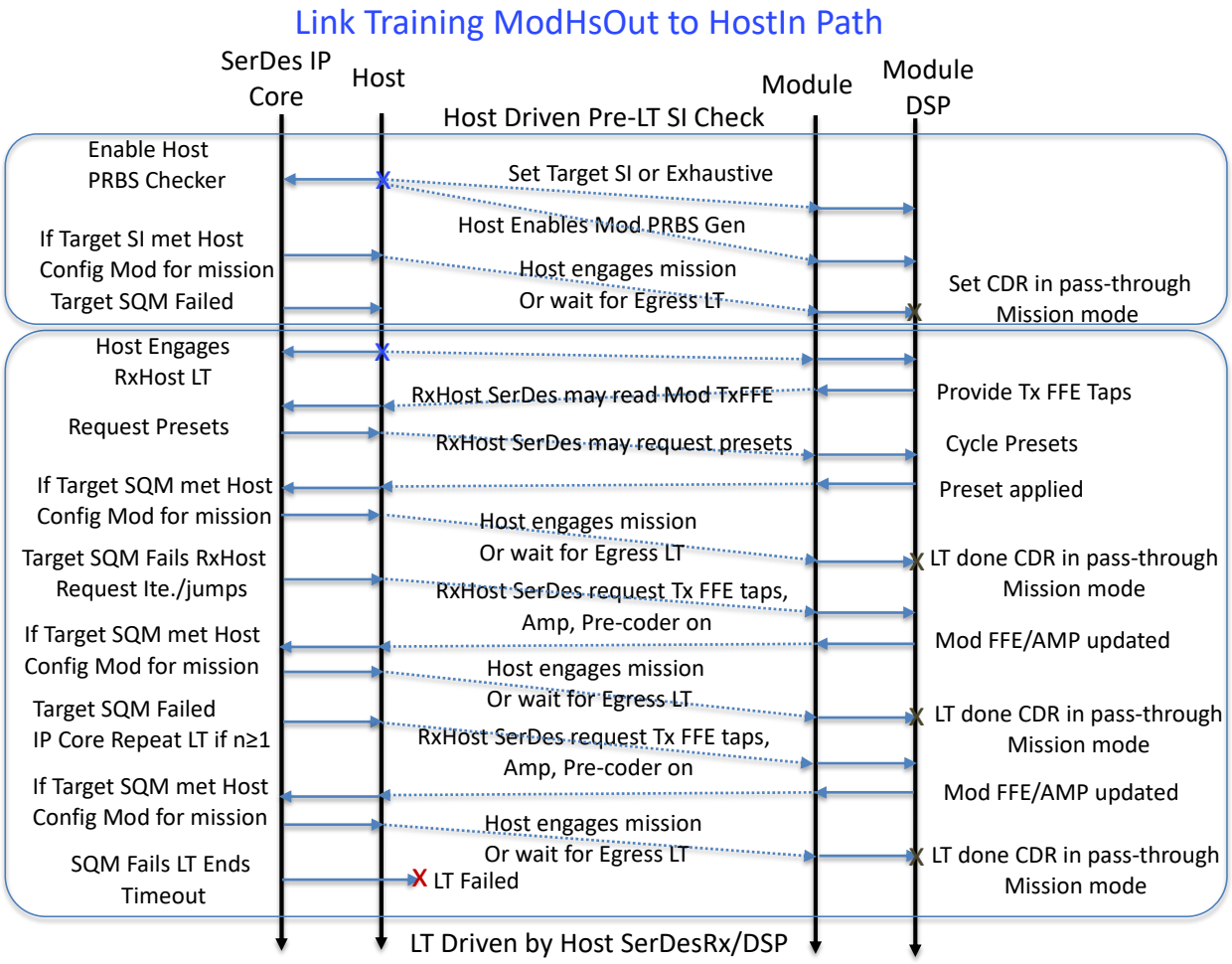

Figure 26: Host SerDes IP Core Initiated LT Transactions

12.2 Module to host path LT operation

Operation of the module LT Engine is very similar to the Host LT operation, except all requests from the module starts with an Interrupt and then host must read and follow with the required action. Interrupt is also used to inform the host if module SerDesRx DSP (RxMod DSP) is trained or if training failed due to timeout.

Key steps for the Module LT are listed below and shown in the flowchart of Figure 27:

I. Host enable RxMod DSP training
   – Module request host to enables HostOut PRBS generator
II. Based on SI parameters SQM, host determines if RxMod need training with default settings
   – If SQM target is met module signal the host then host initiate mission mode
III. Host set target

- Meet best SQM, meet certain SQM

IV. RxMod DSP began link training

- Module requests to the host starts with an interrupt then host read and implements module requests
- Module may request from host to read normalized HostOut FFE coefficients or ModHsIn LT start iterative tuning and adjustment
- Host provide normalized HostOut FFE to the module if requested
- RxMod DSP cylces through preset settings
  - Host informs module HostOut FFE updated
  - if for any preset setting the target SQM is met module informs the host
  - Host may store optimized host HostOut FFE setting
  - Host engages the mission mode
- RxMod DSP tunning continues until the receiver determines it is tuned to its optimal settings or met the target SQM. At that point the settings may be stored to minimize for future link training time.

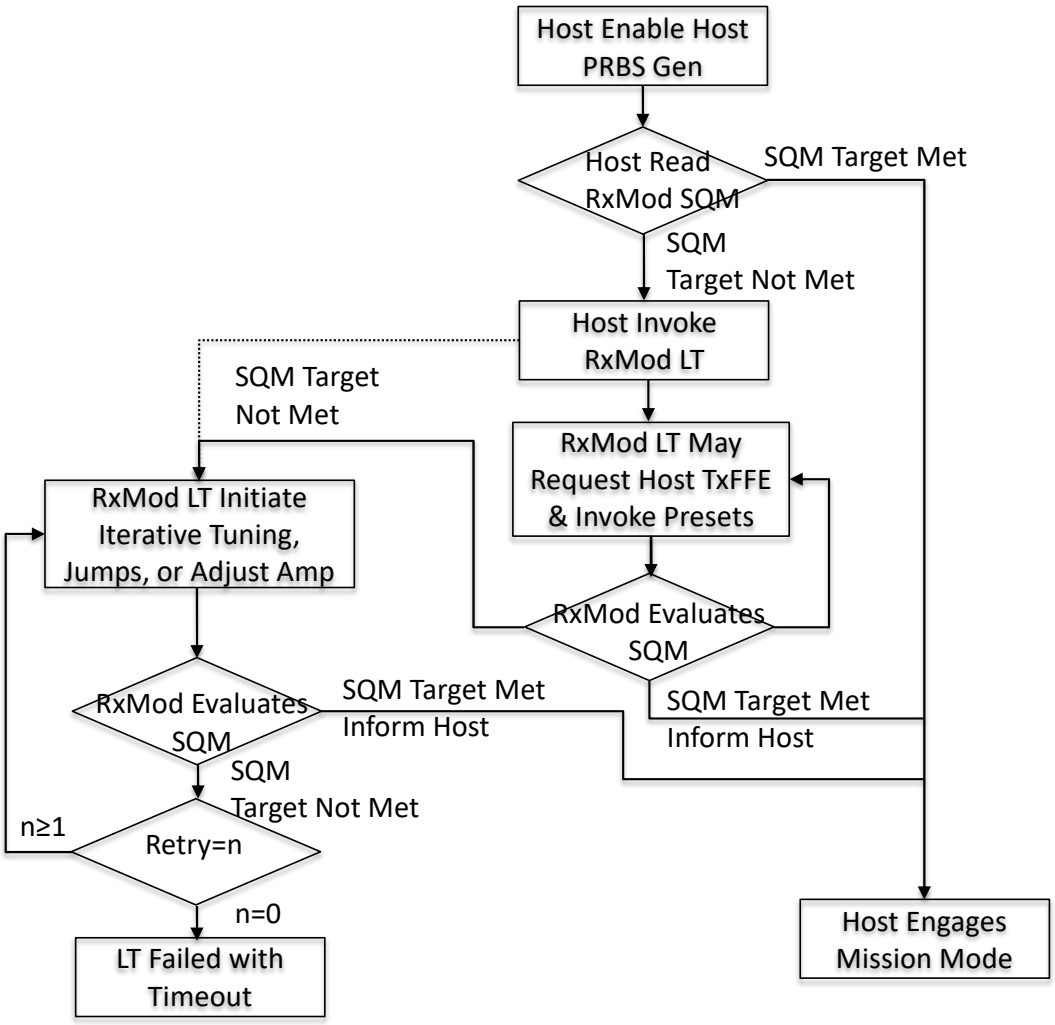

Figure 27: Flowchart of Module to Host LT

Module CDR/DSP initiated LT transactions are shown in Figure 28. The transaction flow diagram includes an optional Pre-LT SI check that relies on prior established baseline SQM to determine if optimum last preset meets the baseline SI, without prior establishes SQM host may read module SQM to determine if link training is needed. If the baseline SerDes SQM is met without link the host may just transition to mission mode indicated by 1st instance of green "X". Exit to mission mode is only possible in the Pre-LT SI check if the host sets reasonable target SQM based on prior LT reported SQM where the optimum preset will meet the target SQM. If the SQM is set to high or set to exhaustive, then exit to mission mode is only possible after 3 instances indicated by green "X" by going through the module CDR/DSP driven LT. If the SQM set to exhaustive where there are no target limits, after one cycle of iterative LT the link will transition to mission mode. Exhaustive LT may be useful to establish the baseline SI or perform full exhaustive LT to establish the best possible SQM. Setting SI parameters to BER and

exhaustive may result in prohibitively longer LT completion time. If link training is not completed within the specified link training timeout, the receive SerDes will declare a link host failure due to timeout.

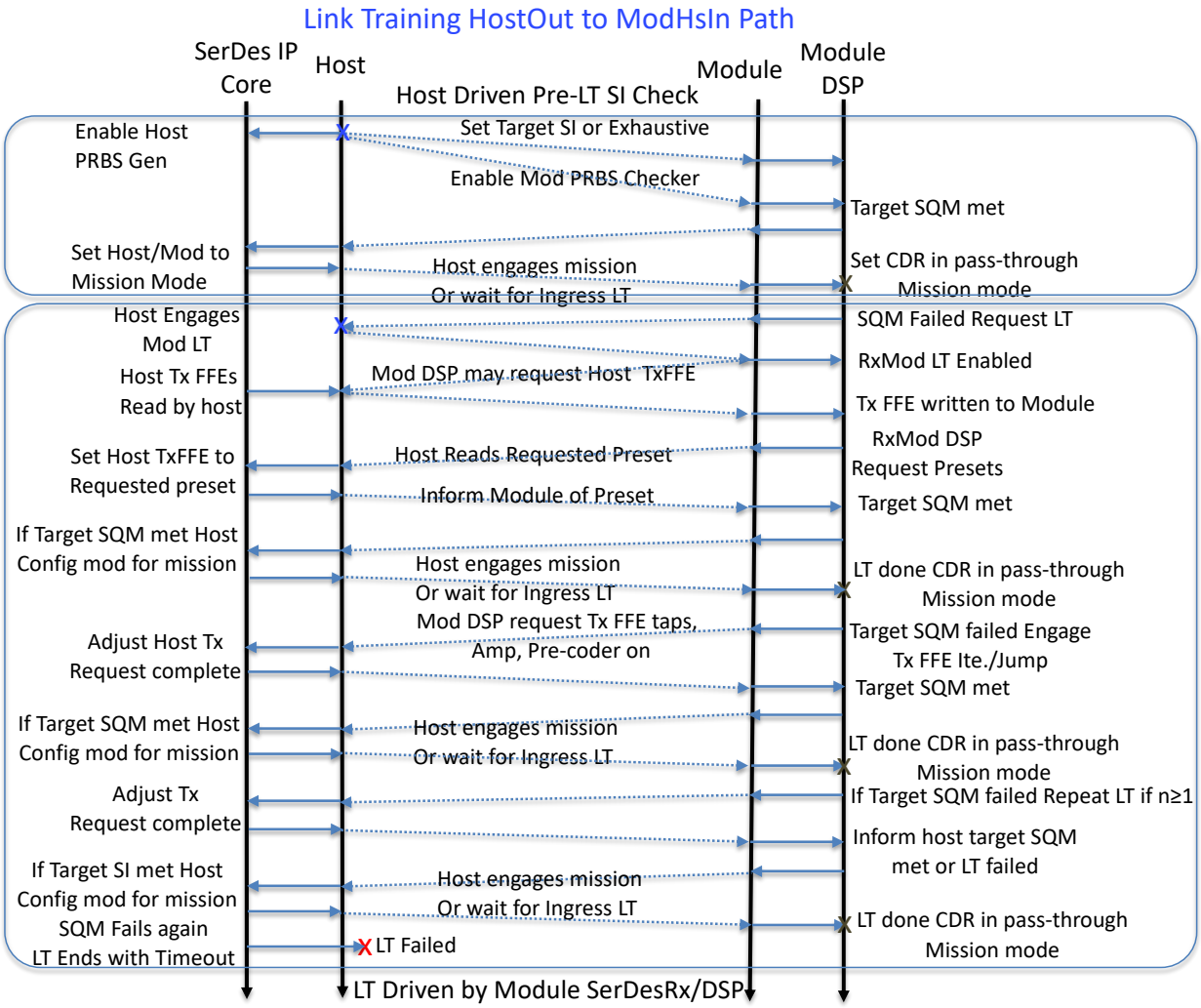

Figure 28: Module CDR/DSP Driven LT Transactions

13 Appendix B - List of References

The following standards and publications are relevant to this specification:

[1] Common Management Interface Specification (CMIS) 5.2, see https://www.oiforum.com/wp-content/uploads/OIF-CMIS-05.2.pdf

[2] IEEE Std 802.3TM-2022, clause 72, clause 92, clause 136, and clause 162.

[3] IEEE Std 802.3ck (100 Gb/s, 200 Gb/s, and 400 Gb/s Electrical Interfaces), clause 120G and 120F.

[4] IEEE 802.3df 200 Gb/s, 400 Gb/s and 800 Gb/s Ethernet Task Force, see https://www.ieee802.org/3/df/index.html.

[5] IEEE 802.3dj 200 Gb/s, 400 Gb/s, 800 Gb/s, and1.6 Tb/s Ethernet Task Force, see https://www.ieee802.org/3/dj/index.html.

[6] OIF CEI-112G-LINEAR-PAM4 proposed draft specification.

[7] OIF Common Electrical I/O CEI 5.1, CL-24 CEI-112G-XSR, CL-25 CEI-112G-VSR, and CL-26 CEI-112G-MR, see https://www.oiforum.com/wp-content/uploads/OIF-CEI-5.1.pdf

[8] OIF-112G-XSR+-PAM4 Extended Reach Extra Short Reach Interface.

[9] Feedforward Equalizer Location Study for High-Speed Serial Systems: Kevin Zheng, Boris Murmann, Hongtao Zhang, and Geoff Zhang - Signal Integrity Journal, May 21, 2019.