

如何实现 PCIe Gen3/Gen4 接收端链路均衡测试

理论篇



目录

术语表.....	3
1 引言.....	4
2 PCIe 3.0 & 4.0 的链路均衡.....	4
2.1 发送端的均衡：FFE.....	5
2.2 接收端的均衡：CTLE 和 DFE.....	7
2.3 链路均衡过程.....	9

术语表

术语	英文	描述
Link		链路。在同一个 LTSSM 状态机的组织下，两个方向上的所有的数据通路组成的整体。
Data Lane		数据通路。一个 Tx 和一个 Rx 通过信号组成的整体。
LEQ	Link Equalization	链路均衡：PCIe 中通过物理层协议的方式，动态地调整链路上双方的均衡设置的过程
AIC	Add-in Card	插卡
SYS	System Board	系统板
DS	Downstream Port	下行端口。例如系统板上的端口、桥器件上面向插卡端的端口均为下行端口。
US	Upstream Port	上行端口。例如插卡上的端口、桥器件上面向系统板端的端口为上行端口。
Preset		预设定值。PCIe 3.0&4.0 中发送端均衡器的一些特定的系数组合。
RC	Root Complex	
EP	Endpoint	

1 引言

PCIe 接口自从被推出以来，已经成为了 PC 和 Server 上最重要的接口。为了更高的数据吞吐率，PCI-SIG 组织于 2010 年推出了 PCIe 3.0，数据速率达到了 8GT/s；于 2017 年推出了 PCIe 4.0，数据速率达到了 16GT/s。PCIe 3.0 已经是 PC 上的标配接口；而 PCIe 4.0 目前仍处于高端 Server 上，在 PC 上也在开始普及。

众所周知，PCIe 1.0 和 2.0 使用的是普通的 FR4 板材和廉价的接插件。为了实现在这样的物理介质上进行有效的信号传输，PCI-SIG 组织使用了 8b/10b 编码和发送端的去加重技术，来保证信号质量：

- 在 PCIe 1.0 中，去加重值为固定值 3.5dB；
- 在 PCI 2.0 中，去加重值为 3.5dB 或 6.0dB 中的任何一个，也是为固定值，无法动态调整。

但是随着速率的翻倍，为了在 PCIe 3.0 和 4.0 中仍然使用普通的 FR4 板材以及廉价的接插件，PCI-SIG 组织对规范做了两方面的改进：

- 使用 128b/130b 编码来代替 8b/10b 编码，使得编码效率大幅提高；
- 使用动态均衡技术，来代替先前代的静态均衡技术；

通过这两方面的改进，PCI-SIG 组织实现了在速率翻倍的同时，仍能保持使用普通的 FR4 板材和廉价接插件。本应用笔记就聚焦于 PCIe 3.0 和 4.0 中的动态均衡技术，介绍其原理、实现及其相关的一致性测试。这样一种动态均衡技术，在 spec 中被称作“Link Equalization”（链路均衡，简称为 LEQ）。

本系列文档包含包含理论篇和实践篇两个部分。理论篇主要介绍 PCIe 3.0/4.0 的链路均衡的工作原理。实践篇则侧重于对链路均衡的测试和调试。

2 PCIe 3.0 & 4.0 的链路均衡

在 PCIe 3.0 和 4.0 中的链路均衡技术相较于先前代要复杂得多，这样一种动态均衡技术可以分为两个方面进行讨论：

- 均衡特性方面：从这个方面来说，相对于先前代的均衡来说，3.0 和 4.0 中的均衡技术的硬件性能指标要求更高了。在本应用笔记中，2.1 节和 2.2 节讨论它的技术细节。
- 协议方面：为了实现动态地调整均衡设置，需要协议层的配合，这是通过 PHY 层的 LTSSM 状态机中的 Recovery.Equalization 子状态来实现的。在本应用笔记中，2.3 节讨论它的技术细节。

我们先来从均衡特性的角度来看 PCIe 3.0 和 4.0 的均衡，如下图 1 展示了在 PCIe 3.0/4.0 中所使用的全部均衡技术，在 Tx 端有 FFE(Feed Forward Equalizer, 前馈均衡器)；在 Rx 端有：CTLE(Continuous Time Linear Equalizer, 连续时间线性均衡器)和 DFE(Decision Feedback Equalizer, 判决反馈均衡器)。通过 FFE 和 CTLE，可以去除大部分由 ISI 所引入的抖动；通过 DFE 可以进一步去除 ISI，它还能去除部分的阻抗失配所造成的反射。通过这些均衡处理，就能够最大程度地保证在接收端判决输入处将眼图打开。

除了上述这些均衡特性上的支持外，在协议层(LTSSM)中还规定需要通过协议的方式来动态调整链路上的均衡设定值，这个过程称作链路均衡(Link Equalization, LEQ)。在链路均衡过程中：

- 本地端按照某个初始 Tx EQ 的设定来发送数据；
- 对端在接收到数据时，会根据误码率或信号质量来判断该 Tx EQ 是否合适；
- 若不合适，对端会通过协议向本地端请求一个新的 Tx EQ 值；
- 本地端在接收到这个请求值之后，会改变 Tx EQ 的值。

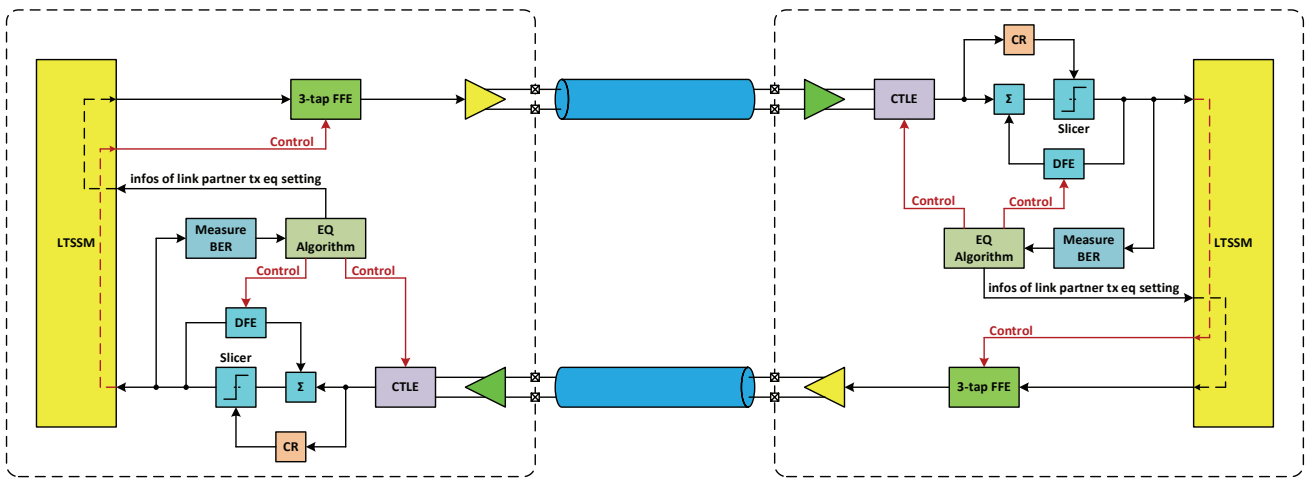
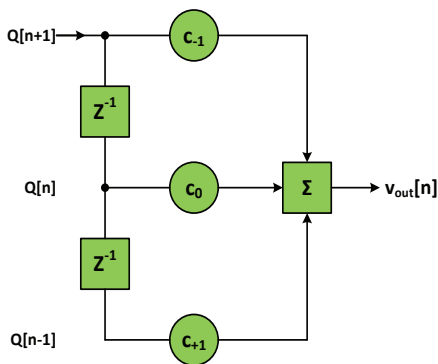


图 1. LEQ 硬件实现的模块框图

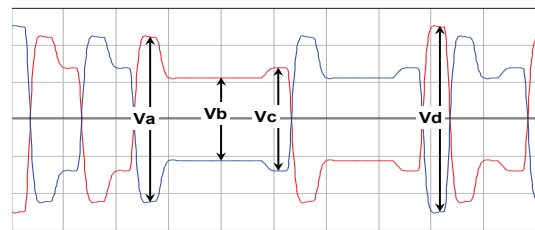
通过这一动态过程，就能够保证链路上的 Tx EQ 为最优值。与此同时，本地端和对端也会同时调整 Rx EQ。通过动态地调整 Tx EQ 和 Rx EQ，就能够灵活地适应不同的信道情况。

2.1 发送端的均衡：FFE

在 PCIe 3.0 & 4.0 中使用的都是 3-tap FFE，如图 2a 所示。其中， $Q[n]$ 为数字信号，建模时取值为 ± 1 ； $\{c_{-1}, c_0, c_{+1}\}$ 为 FFE 的抽头系数； $v_{out}[n]$ 为发送端的模拟信号输出。为了后续讨论的方便，会对系数进行归一化处理，如式 (2) 所示，这样就会在式 (1) 中多了一个比例系数 g 。



(a) FFE 的模型框图



$$\begin{aligned} \text{De-emphasis} &= 20 \log_{10} Vb/Va \\ \text{Preshoot} &= 20 \log_{10} Vc/Vb \\ \text{Boost} &= 20 \log_{10} Vd/Vb \end{aligned}$$

(b) FFE 的模拟电压输出

图 2. PCIe 3.0 & 4.0 发送端所使用的 3-tap 的 FFE

$$v_{out}[n] = (c_{-1}Q[n + 1] + c_0Q[n] + c_{+1}Q[n - 1])g \quad (1)$$

$$|c_{-1}| + |c_0| + |c_{+1}| = 1, c_{-1} \leq 0, c_{+1} \leq 0 \quad (2)$$

根据式 (1) 不难看出, 理想情况下的差分电压幅度有: $2^3 \div 2=4$ 种可能性, 这四个电压幅度在 PCIe 标准中(如图 2b 所示) 分别被标记为 Va, Vb, Vc, Vd:

$$\begin{aligned} V_a &= (c_0 - c_{+1} + c_{-1}) \times 2g \\ V_b &= (c_0 + c_{+1} + c_{-1}) \times 2g \\ V_c &= (c_0 + c_{+1} - c_{-1}) \times 2g \\ V_d &= (c_0 - c_{+1} - c_{-1}) \times 2g \end{aligned}$$

其中, Vb 被称作去加重电压 (de-emphasis voltage), Vc 被称作预冲电压 (preshoot voltage); Vd 被称作最大幅度电压 (boost voltage), PCIe 标准中没有为 Va 取一个专门的名字。在此基础上, 标准中通过三组比值来完备地描述 FFE 的性能:

$$DEEMPHASIS = 20 \log \frac{V_b}{V_a} = 20 \log \frac{c_0 + c_{+1} + c_{-1}}{c_0 - c_{+1} + c_{-1}} \quad (3)$$

$$PRESHOOT = 20 \log \frac{V_c}{V_b} = 20 \log \frac{c_0 + c_{+1} - c_{-1}}{c_0 + c_{+1} + c_{-1}} \quad (4)$$

$$BOOST = 20 \log \frac{V_d}{V_b} = 20 \log \frac{c_0 - c_{+1} - c_{-1}}{c_0 + c_{+1} + c_{-1}} \quad (5)$$

若不加限制的来说, 那么 $\{c_{-1}, c_0, c_{+1}\}$ 形成的组合有无穷多个。但并不是所有的组合在实际应用中都是合适的。其中一个最重要的约束条件就是: 去加重电压 Vb 不能过小, 过小的去加重电压会导致输出信号在接收端的眼高过低。因此通过 BOOST 比值对去加重地电压幅值进行限制: 对于满摆幅的 Tx 输出, 规范要求 BOOST ≤ 9.5 dB; 对于减摆幅的 Tx 输出, 规范要求 BOOST ≤ 3.5 dB。最终会形成一个如图 3 类似的矩阵表, 图中系数的粒度为 1/24。在实际应用中可以是其他的粒度值, 例如 1/64; 更小的粒度能够使系数空间的取值可能性更多, 在 LEQ 调节时也更精细。

Min Reduced Swing Limit

PS	DE	C_{+1}								
		0/24	1/24	2/24	3/24	4/24	5/24	6/24	7/24	8/24
C_{-1}	BOOST	0.0 0.0	0.0 -0.8	0.0 -1.6	0.0 -2.5	0.0 -3.5	0.0 -4.7	0.0 -6.0	0.0 -7.6	0.0 -9.5
	0/24	0.0 0.0	0.8 0.8	1.6 1.6	2.5 2.5	3.5 3.5	4.7 4.7	6.0 6.0	7.6 7.6	9.5 9.5
	1/24	0.8 0.8	0.8 -0.8	0.9 -1.7	1.0 -2.8	1.2 -3.9	1.3 -5.3	1.6 -6.8	1.9 -8.8	2.5 -10.5
	2/24	1.6 1.6	1.7 -0.9	1.9 -1.9	2.2 -3.1	2.5 -4.4	2.9 -6.0	3.5 -8.0	4.1 -10.5	4.7 -13.0
	3/24	2.5 2.5	2.8 -1.0	3.1 -2.2	3.5 -3.5	4.1 -5.1	4.9 -7.0	5.7 -9.5	6.5 -12.0	7.3 -15.0
	4/24	3.5 3.5	3.9 -1.2	4.4 -2.5	5.1 -4.1	6.0 -6.0	7.6 -8.5	9.5 -11.0	11.5 -14.0	13.5 -17.0
	5/24	4.7 4.7	5.3 -1.3	6.0 -2.9	7.0 -4.9	8.5 -7.0	10.5 -10.0	13.0 -13.0	16.0 -16.0	19.5 -19.5
6/24	6.0 6.0	6.8 -1.6	8.0 -3.5	9.5 -6.0	11.5 -9.5	14.5 -13.5	18.0 -17.0	22.0 -21.0	26.5 -25.5	

Full swing Limit or Max reduced swing limit

图 3. 发送端均衡的系数空间的矩阵表举例

鉴于系数空间上的取值可能性较多，PCI-SIG 协会在开发协议的过程中，广泛地研究了在不同插入损耗下最优的系数取值组合；最后选定了若干个特定的系数取值组合，并把它们称作预设值（preset），如表 1 所示。这样，在实际的 LEQ 过程中，链路双方就可以先采用预设值进行粗调；若还认为链路的均衡设置仍然没有达到最优，可以进一步通过系数空间的方

式进行细调。显然，粗调速度较快，但跨度可能比较大；细调则速度较慢，而调节较精细；这样通过先粗调，再细调，就能够达到速度和精度的平衡。选择使用预设值或使用系数，是通过 TS1 序列中 Symbol 6 bit[7] 字段的取值来实现的；设定为 1b 时，使用预设值进行 LEQ 的调节；设定为 0b 时，使用系数组合进行 LEQ 的调节。

表 1. PCIe 3.0&4.0 中的预设值

Preset Number	Preshoot (dB)	De-emphasis (dB)	c-1	c+1	Va/Vd	Vb/Vd	Vc/Vd
P4	0.0	0.0	0.000	0.000	1.000	1.000	1.000
P1	0.0	-3.5 ± 1dB	0.000	-0.167	1.000	0.668	0.668
P0	0.0	-6.0 ± 1.5dB	0.000	-0.250	1.000	0.500	0.500
P9	3.5 ± 1dB	0.0	-0.166	0.000	0.668	0.668	1.000
P8	3.5 ± 1dB	-3.5 ± 1dB	-0.125	-0.125	0.750	0.500	0.750
P7	3.5 ± 1dB	-6.0 ± 1.5dB	-0.100	-0.200	0.800	0.400	0.600
P5	1.9 ± 1dB	0.0	-0.100	0.000	0.800	0.800	1.000
P6	2.5 ± 1dB	0.0	-0.125	0.000	0.750	0.750	1.000
P3	0.0	-2.5 ± 1dB	0.000	-0.125	1.000	0.750	0.750
P2	0.0	-4.4 ± 1.5dB	0.000	-0.200	1.000	0.600	0.600
P10	0.0	var	0.000	var	1.000	var	var

2.2 接收端的均衡：CTLE 和 DFE

在 PCIe 3.0 & 4.0 Base 规范中，并没有明确地规定接收端的结构是怎样的；而只是从测量的角度对接收端性能进行了规定。相反地，在规范中定义了一个行为级 CTLE 和行为级 DFE。这些行为级模型可以作为设计指南；并且为了使得待测对象能够通过规范的要求，一般来说用户所设计的接收端性能至少要等于这

些行为级模型的性能，可以强于这些行为级模型，但不能弱于这些行为级模型。PCIe 3.0 & 4.0 的行为级 CTLE 的传递函数可以表示为：

$$H(s) = \omega_{P2} \frac{s + A_{DC} \omega_{P1}}{(s + \omega_{P1})(s + \omega_{P2})} \quad (6)$$

其中， A_{DC} 为直流增益， ω_{P1} 为低频极点； ω_{P2} 为高频极点，各参数的取值见表 2 所示。

表 2. 行为级 CTLE 的参数取值

参数	8 GT/s	16 GT/s
A_{DC}	-6.0 dB ~ -12.0 dB, -1.0 dB step	
ω_{P1}	$2\pi \times 2$ GHz	$2\pi \times 2$ GHz
ω_{P2}	$2\pi \times 8$ GHz	$2\pi \times 16$ GHz

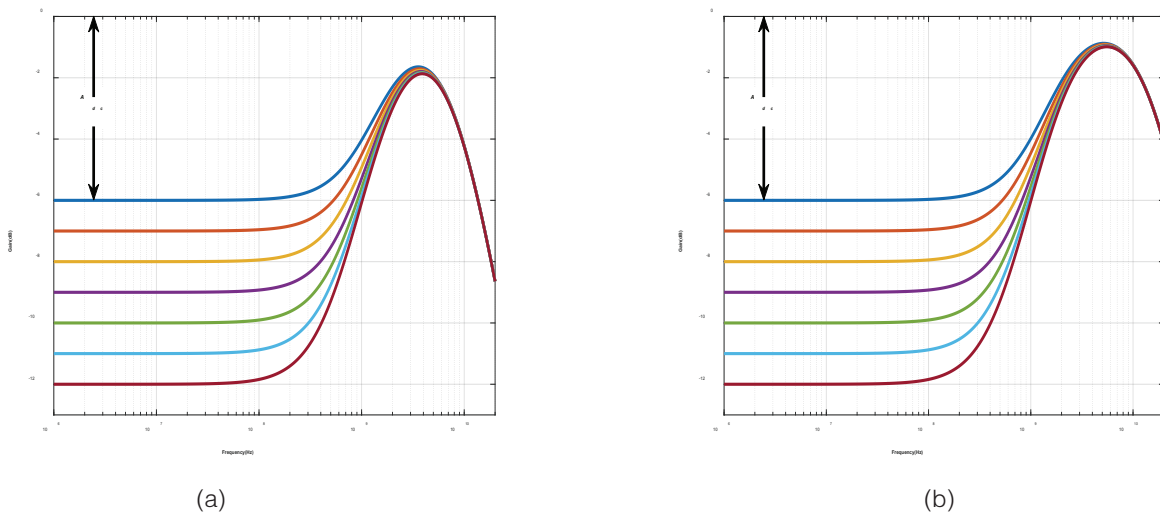


图 4. 行为级 CTLE 的频响曲线：(a) PCIe 3.0 (b) PCIe 4.0

发送端的输出在经过一段很长的 FR4 走线之后，仅仅使用 CTLE，可能是不够的。因此在 PCIe 3.0 & 4.0 中，还使用了 DFE 的技术。在 3.0 中，使用 1-tap 的 DFE，而在 4.0，由于速率相对于 3.0 翻倍了；所以使用 2-tap 的 DFE，以便移除更大的 ISI。

与线性均衡器 FFE 和 CTLE 相比，DFE 为一种非线性均衡器。DFE 的基本想法是：若已经正确接收了之前的比特数据的话；那么先前的比特数据对当前比特

所产生的影响就是已知的；从而我们就可以通过反馈的方式进行补偿，这样就能够进一步消除抖动和噪声的影响。不难看出这里的非线性体现在：反馈回来的信号是经过判决之后的数字信号；而判决电路是一种非线性电路。显然，反馈通路上的抽头数目越多，那么对抖动和噪声的消除可能就越好；这也就是为什么 3.0 中使用 1-tap 的 DFE，而在 4.0 中使用 2-tap 的 DFE。

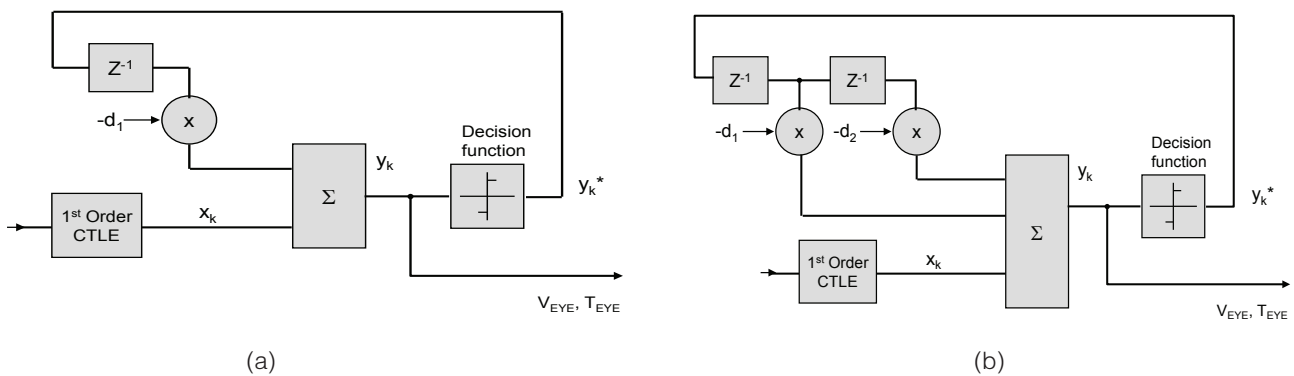


图 5. 行为级 DFE 的结构：(a) PCIe 3.0 (b) PCIe 4.0

2.3 链路均衡过程

在 2.1 节和 2.2 节中，我们主要从硬件特性的角度考察了 PCIe 3.0 & 4.0 所使用的均衡技术。但是这个远远是不够的。

链路的两端刚开始建立通信的时候，并不知道整个信道的物理特性是怎样的，例如插入损耗多大，是否有阻抗不连续等。由于 PCIe 3.0 和 4.0 的插入损耗允许的变化范围很大，一个静态的均衡设置并不能覆盖所有的情况。这样就需要链路的两端根据当前物理信道的特性，来动态地调整均衡设置，使得均衡设置对于当前的物理信道来说是最优的。假设 Port A 和

Port B 是一个链路上的两端，那么链路均衡过程要做的事情有：

- 配置 Port A 和 Port B 的初始均衡设置；
- 配置从 Port A Tx → Port B Rx 这一方向的均衡设置；
- 配置从 Port B Tx → Port A Rx 这一方向的均衡设置；

下面我们以 Port A Tx → Port B Rx 这一方向来说明链路均衡时如何实现的。如图 6 所示，在 8GT/s 或者 16GT/s 速率下的链路开始建立通信时，是以初始的未优化的 TX EQ 在发送 TS1/TS2 序列，并且 Port A 在 TS1/TS2 序列中表明其所用的 TX EQ 的值。

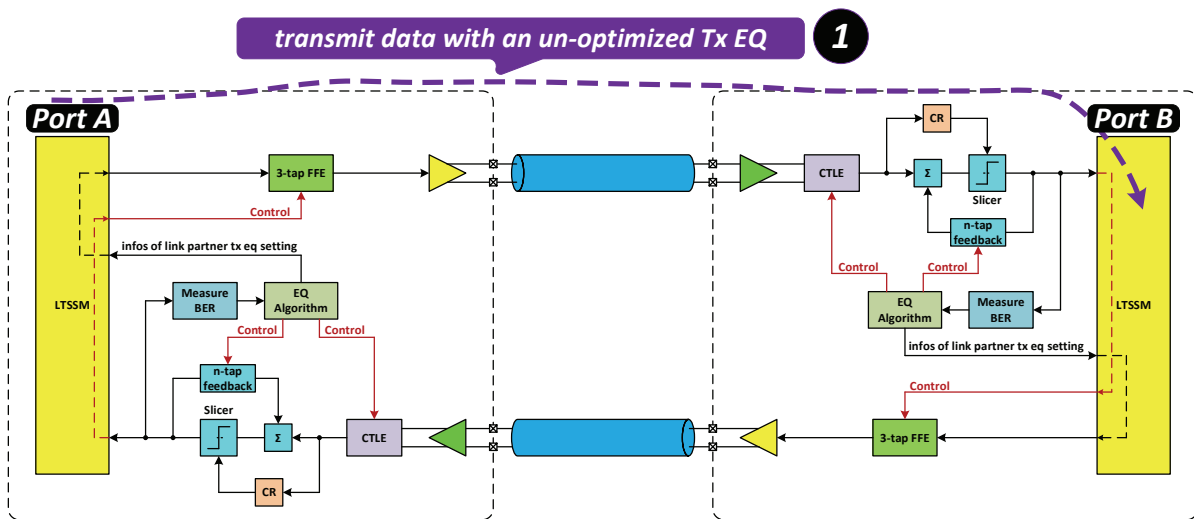


图 6. LEQ: 本地端发送未经优化的初始 TX EQ

当 Port B Rx 在接收到这些 TS1/TS2 序列时，芯片内部存在一块电路或者一套算法来评估当前的 TX EQ 是否合适，若认为不合适，就会如图 7 所示，发送 TS1 序列来请求一个新的 TX EQ。

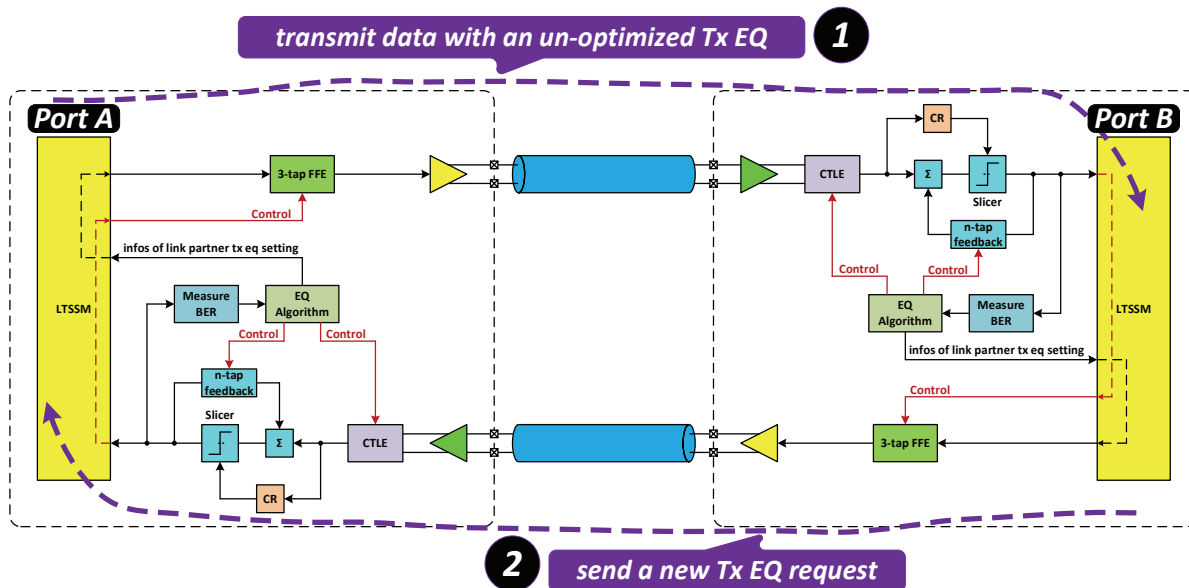


图 7. LEQ: 对端请求一个新的 TX EQ

随后，Port A 会接收到请求设置 TX EQ 的 TS1 序列，如图 8 所示，调整其 TX 端的 FFE 的设置。

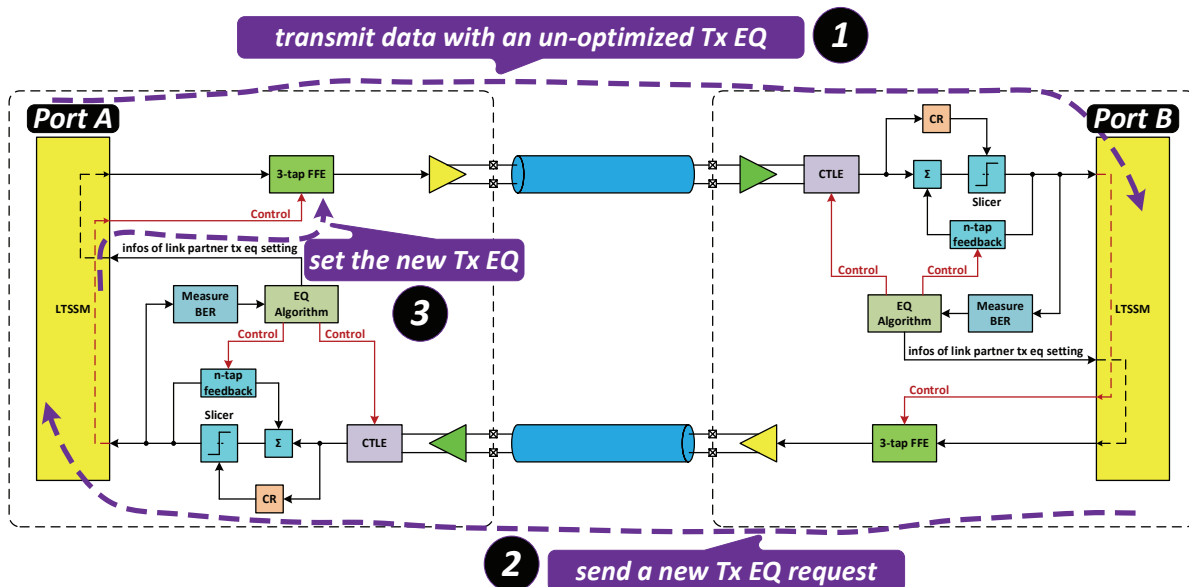


图 8. LEQ: 本地正确地接收到了对端的请求，设置新的 TX EQ

Port A 在调整完 Tx FFE 的设置之后，如图 9 所示，会将新的 TX EQ 设置值更新到 TS1/TS2 的序列之中，发送到 Port B 端。若 Port B 仍然觉得这个时候的 TX

EQ 不是最优，那么仍然会重复图中的 2~4 步骤，直到达到最优的 TX EQ。当然上述过程并不能无限进行下去，必须要在大概 32ms 的范围进行完。

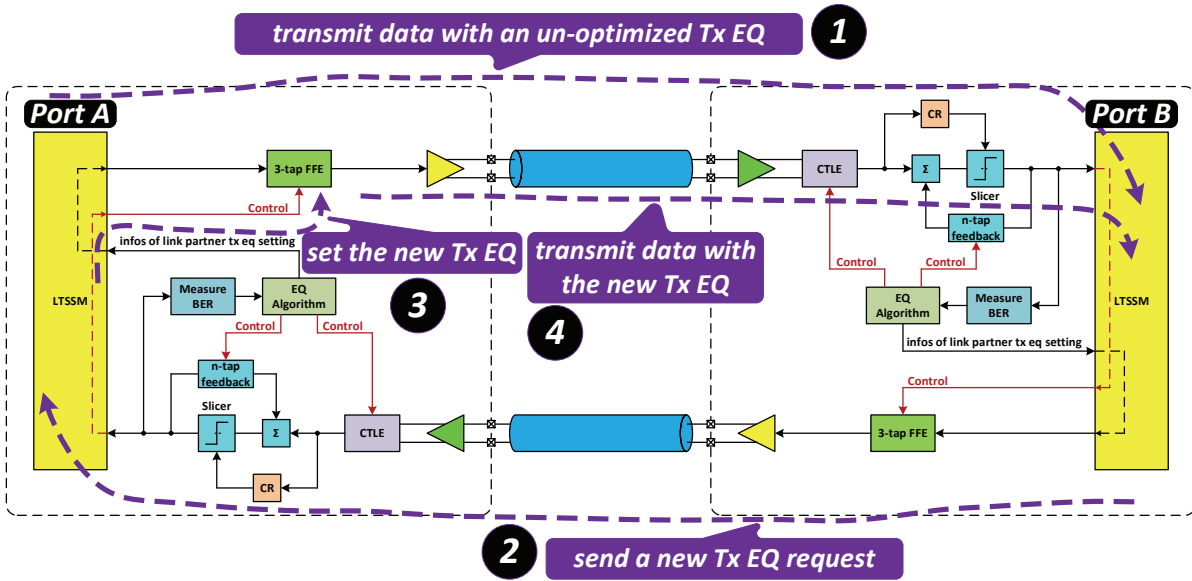


图 9. LEQ: 本地端告知对端已成功设置新的 TX EQ

在上述 2~4 步骤的同时，Port B 的 RX 端也在不停地调整其 RX EQ，如图 10 所示。

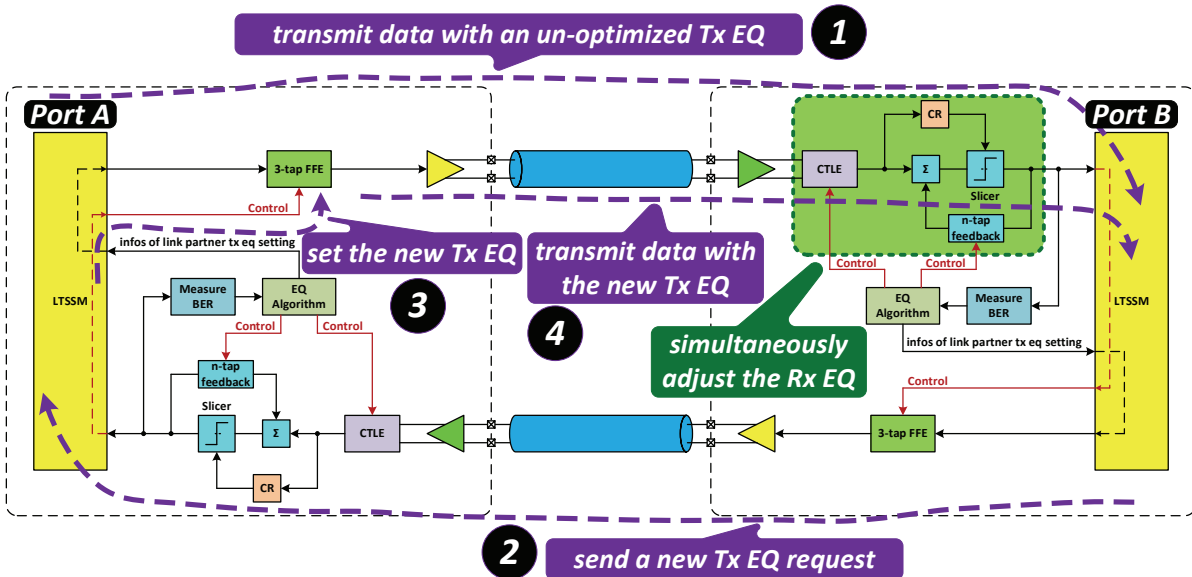


图 10. LEQ: 整个过程中同时调整 RX EQ

Port B Tx → Port A Rx 这一方向的链路均衡过程大体上是一样的。但是会有细微的不同，这一不同是由于一个链路上的两个端口的角色不对等所造成的。

众所周知，PCIe 系统中的端口，存在两种角色：下行端口（Downstream Port）和上行端口（Upstream Port）。比如系统板上的 RC（Root-Complex）就是下行端口；而插卡上的 EP（Endpoint）为上行端口；而 PCIe 桥器件上，面向 RC 端的为上行端口，面向 EP 端的为下行端口。

下行端口和上行端口的角色不对等在 LEQ 中的体现为：下行端口的初始 TX EQ 值由其配置寄存器所直接指定；而上行端口的初始 TX EQ 则不是由自身决定的，而是在 LEQ 之前由其对端的下行端口通过 TS1 序列传送给它的。

如图 6~ 图 10 中所讨论的，LEQ 是基于请求 - 响应机制来完成动态均衡的。在 PCIe 的规范中，LEQ 总共包含四个阶段：Phase 0、Phase 1、Phase 2、Phase 3。其中上行端口包含全部四个过程；而下行端口不包含 Phase 0。通过图 11 不难看出，在 LEQ 过程中，上行端口和下行端口的行为是有区别的。每一个阶段的动作如下所示：

- **Phase 0:** 仅仅只有上行端口才有该状态。上行端口的 Tx EQ 设定值由下行端口在 Recovery. Rcvrcfg 状态下通过 TS1 传送给上行端口。在 Phase 0 阶段，上行端口按照这个初始 Tx EQ 来发送 TS1 序列。

- **Phase 1:** 对于下行端口来说，其 Tx EQ 的初始值来自其自身的配置寄存器。因此无论是下行端口还是上行端口，在该阶段均发送初始 Tx EQ 的 TS1 序列。此时对该链路（Link）上两个方向的数据通路（Data Lane）的误码率要求为： $BER < 1e-4$ 。
- **Phase 2:** 在这个阶段，上行端口通过协议请求的方式对下行端口的 Tx EQ 进行优化，以便使得该方向的数据通路（Data Lane）的误码率在该阶段完成之后满足： $BER < 1e-12$ 。此时上行端口为请求方；而下行端口为响应方。
- **Phase 3:** 此阶段的过程与 Phase 2 类似，只不过角色发生翻转。在这个阶段，下行端口通过协议请求的方式对上行端口的 Tx EQ 进行优化，以便使得该方向的数据通路（Data Lane）的误码率在该阶段完成之后满足： $BER < 1e-12$ 。此时下行端口为请求方；而上行端口为响应方。

以上描述的是在 LEQ 过程中链路上的双方如何调整 Tx EQ。而对于 Rx EQ，根据 Base 规范中的说明，在整个 LEQ 的过程、以及在后续正常工作的过程中，链路双方都可以一直调整 Rx EQ。

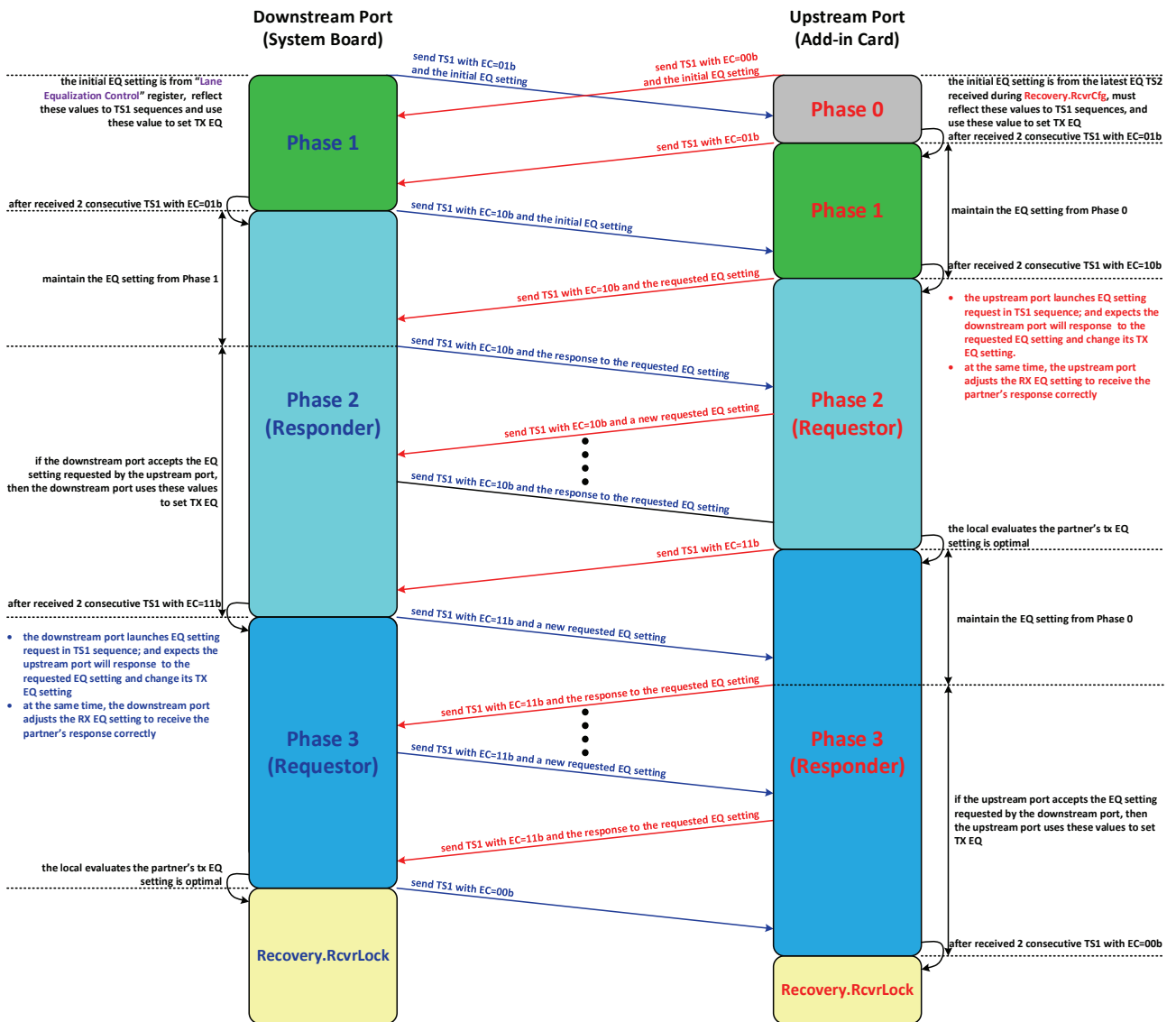


图 11. LEQ 的状态跳转示意图

泰克科技(中国)有限公司

上海市浦东新区川桥路1227号
邮编: 201206
电话: (86 21) 5031 2000
传真: (86 21) 5899 3156

泰克北京办事处

北京市朝阳区酒仙桥路6号院
电子城·国际电子总部二期
七号楼2层203单元
邮编: 100015
电话: (86 10) 5795 0700
传真: (86 10) 6235 1236

泰克上海办事处

上海市长宁区福泉北路518号
9座5楼
邮编: 200335
电话: (86 21) 3397 0800
传真: (86 21) 6289 7267

泰克深圳办事处

深圳市深南东路5002号
信兴广场地王商业大厦3001-3002室
邮编: 518008
电话: (86 755) 8246 0909
传真: (86 755) 8246 1539

泰克成都办事处

成都市锦江区三色路38号
博瑞创意成都B座1604
邮编: 610063
电话: (86 28) 6530 4900
传真: (86 28) 8527 0053

泰克西安办事处

西安市二环南路西段88号
老三届世纪星大厦26层L座
邮编: 710065
电话: (86 29) 8723 1794
传真: (86 29) 8721 8549

泰克武汉办事处

武汉市洪山区珞喻路726号
华美达大酒店702室
邮编: 430074
电话: (86 27) 8781 2760

泰克香港办事处

香港九龙尖沙咀弥敦道132号
美丽华大厦808-809室
电话: (852) 2585 6688
传真: (852) 2598 6260

如需更多资源, 敬请访问 WWW.TEK.COM.CN。

© 泰克科技公司版权所有, 侵权必究。泰克产品受到已经签发及正在申请的美国专利和国外专利保护。本文中的信息代替所有以前出版的材料中的信息。技术数据和价格如有变更, 恕不另行通告。TEKTRONIX 和泰克徽标是泰克公司的注册商标。本文提到的所有其他商号均为各自公司的服务标志、商标或注册商标。

2020年2月

