



# PCI Express® Link Training and Protocol Debug Techniques

**Matthew Dunn**  
**Global Technical Services Manager**  
**Teledyne LeCroy**

# Disclaimer



**Presentation Disclaimer: All opinions, judgments, recommendations, etc. that are presented herein are the opinions of the presenter of the material and do not necessarily reflect the opinions of the PCI-SIG®.**

# Objective



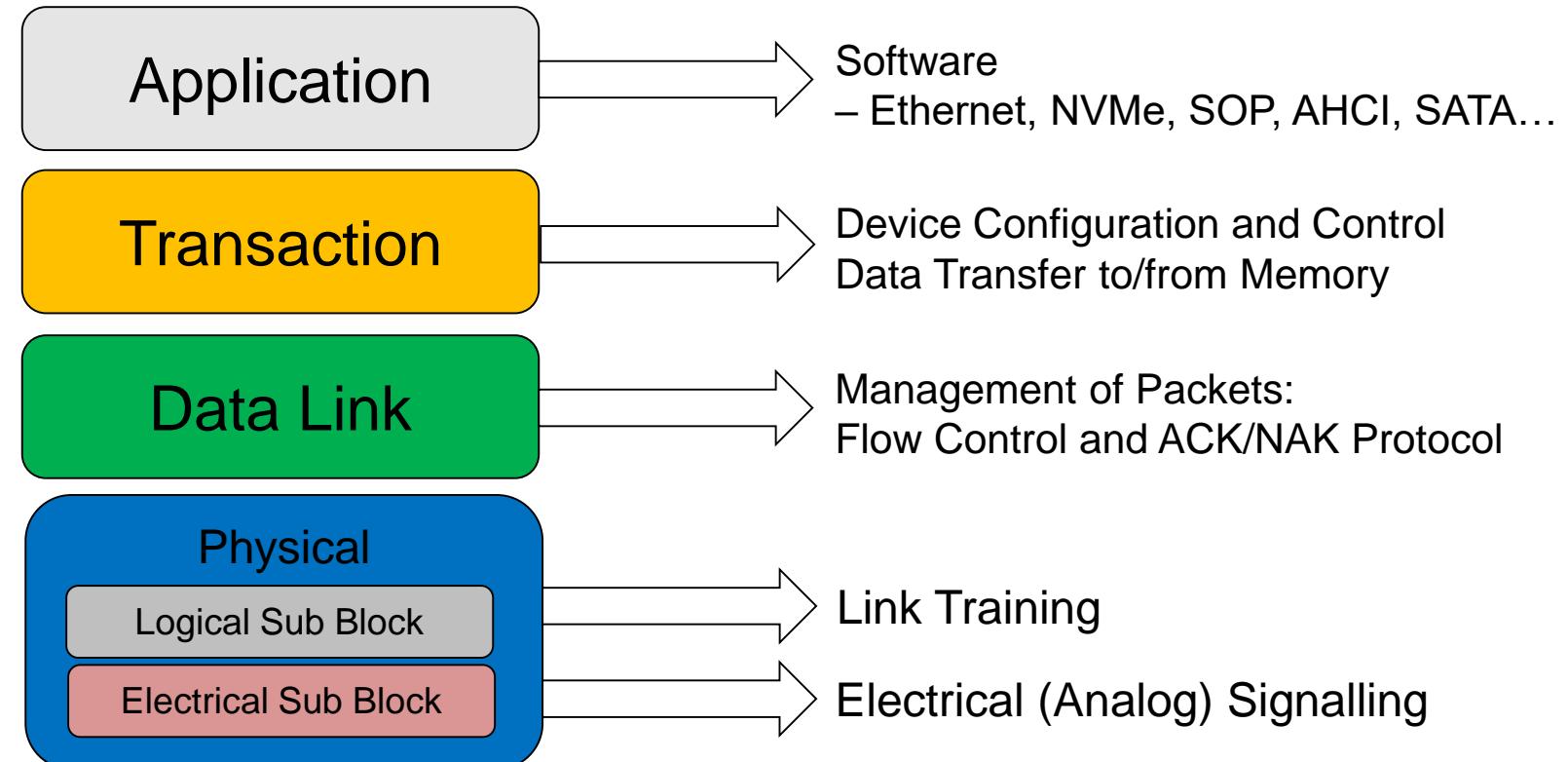
- The examples are all real problems that have been observed multiple times.
- This presentation will discuss techniques applicable to testing Link Training and Protocol for all PCI Express® devices, including those designed to the PCI Express 4.0 Specification and beyond.

# Agenda

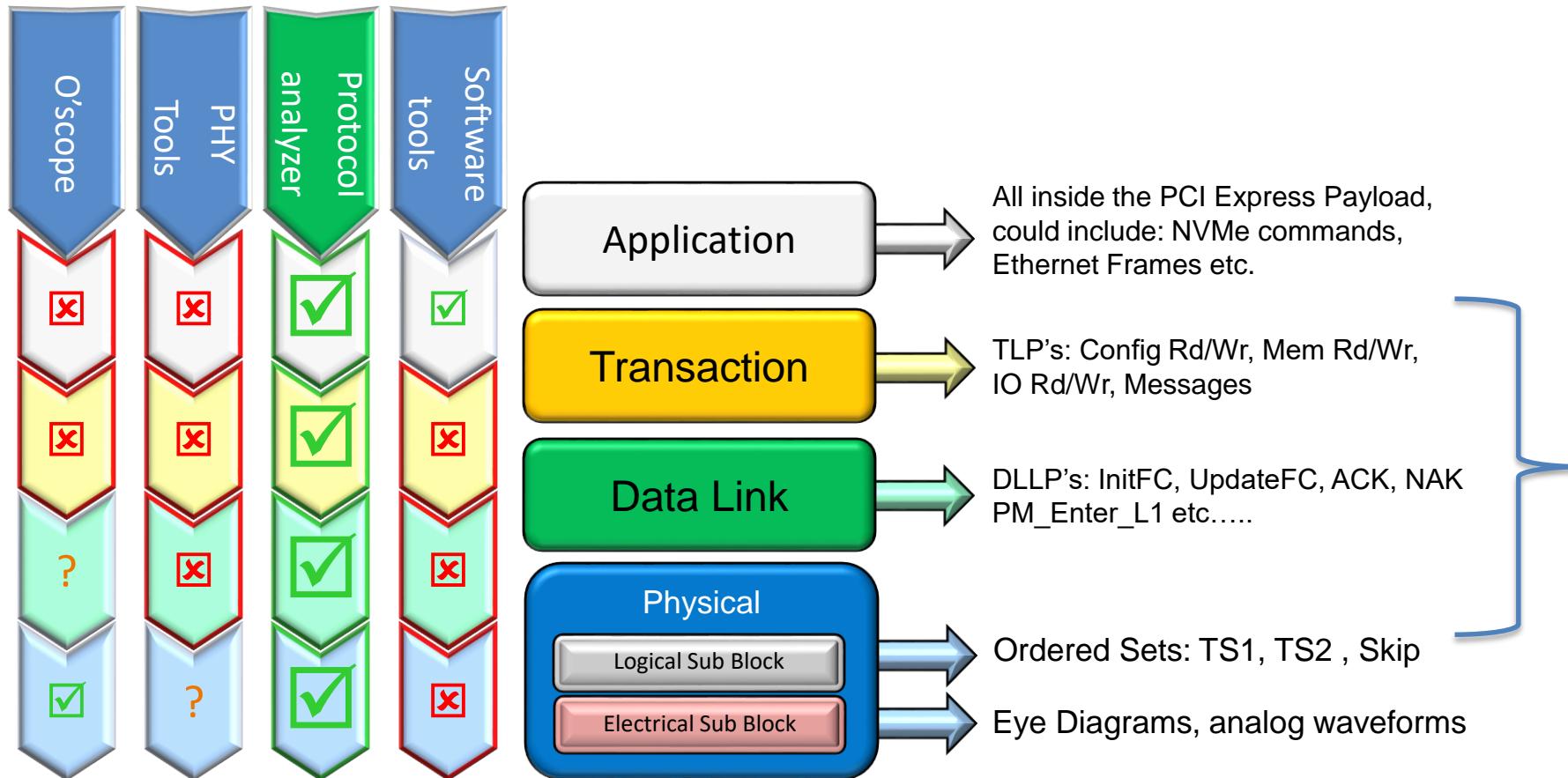


- **PCI Express Layered Model Overview**
- **Top PCIe® Protocol Problems**
  - Link Training
  - Polarity
  - Lane Reversal
  - Power Management
  - Reset Issues
  - Flow Control Problems
  - Class Code
  - Compliance Test Failures
  - Device ID Changes

# PCI Express® Layered Model



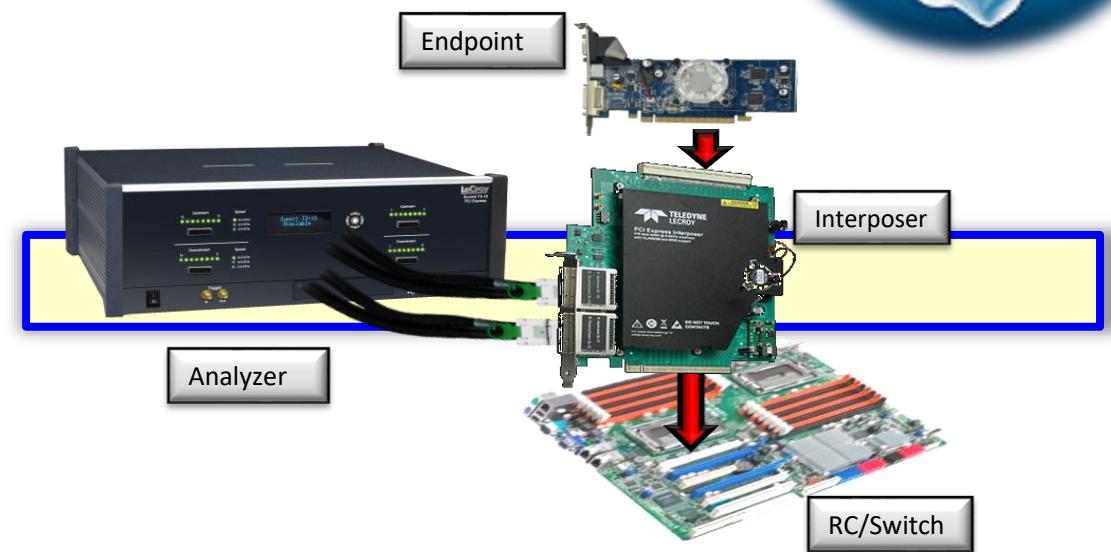
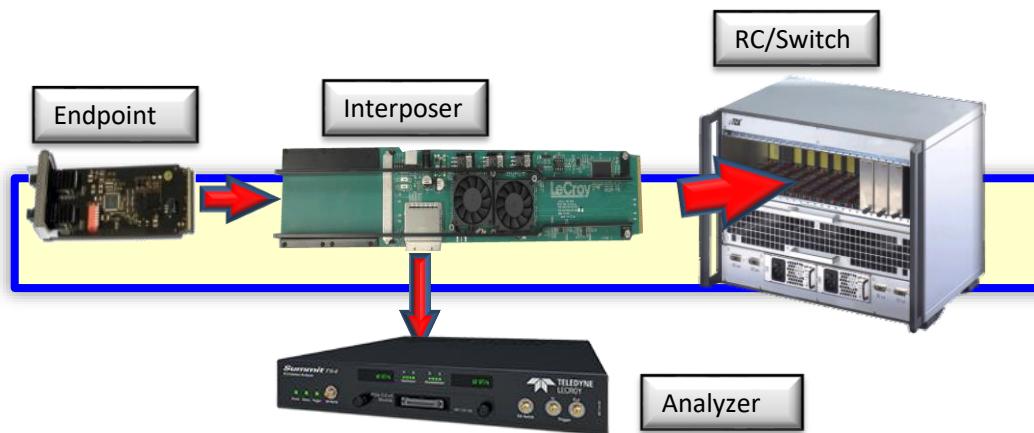
# If I Had to Choose One Tool...



# Protocol Analyzers See Below the Waterline

- **Protocol Analyzer**

- Visibility of layers not exposed by Software or Physical layers
- Simplify signals to byte/dWord
- Allow filtering of scrambled data
- Extend recording times
- Real time triggers



# What Can I See at Each Layer?

## Application

NVM Cmd	H	OPC	SQID	CQID	PRP1	QID	QSIZE	PC	QPRIo	CQID	ST	SC	SCT	Device ID	NSID	Metrics	# NVME Trans	Time Delta	Time Stamp
12		Create I/O SQ	0x0000	0x0000	00000001:80F30000	0x0004	0x03FF	1	0x2	0x0004	ST	Successful Completion	Generic Command Status	001:00	0x00000000	4	1.000 ms	0013 . 772 658 950 s	
NVM Cmd	H	OPC	SQID	CQID	PRP1	PRP2		FID	SEL	TMPTH	ST	SC	SCT	Device ID	NSID	Metrics	# NVME Trans	Time Delta	
13		Get Features	0x0000	0x0000	00000000:00000000	00000000	00000000	Temperature Threshold	0x0	0x0000	ST	Successful Completion	Generic Command Status	001:00	0x00000000	4	1.000 ms		
NVM Cmd	H	OPC	SQID	CQID	PRP1	PRP2		FID	SV	SMART	ST	SC	SCT	Device ID	NSID	Metrics	# NVME Trans	Time Delta	
14		Set Features	0x0000	0x0000	00000000:00000000	00000000	00000000	Asynchronous Event Configuration	0	0x00	ST	Successful Completion	Generic Command Status	001:00	0x00000000	4	1.000 ms		

## Transaction

Split Tra	R→	8.0	Cfg	CfgRd0	RequesterID	CompleterID	Tag	TC	VC ID	DeviceID	Register	Status	Expansion ROM Base Address	Metrics	# LinkTrans	Time Delta	Time Stamp	
894760		x4		000:00100	000:00:0	001:00:0	0	0	001:00:0	0x030	SC	0xF7E00000	2	2	2.695 us	0007 . 491 028 842 s		
Split Tra	R→	8.0	Cfg	CfgWr0	RequesterID	CompleterID	Tag	TC	VC ID	DeviceID	Register	Status	Expansion ROM Base Address	Metrics	# LinkTrans	Time Delta	Time Stamp	
894761		x4		010:00100	000:00:0	001:00:0	0	0	001:00:0	0x030	SC	0xF7E00001	2	2	4.266 us	0007 . 491 031 536 s		
Split Tra	R→	8.0	Cfg	CfgWr0	RequesterID	CompleterID	Tag	TC	VC ID	DeviceID	Register	Status	Status	Command	Metrics	# LinkTrans	Time Delta	Time Stamp
894762		x4		010:00100	000:00:0	001:00:0	0	0	001:00:0	0x004	SC	0x0000	0x0002	2	2	5.116 us	0007 . 491 035 802 s	
Split Tra	R→	8.0	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data	Data	Metrics	# LinkTrans	Time Delta	Time Stamp	
894763		x4		000:00000	000:00:0	001:00:0	0	0	F7E00000	SC	1 dword	1 dword	2	2	4.190 us	0007 . 491 040 918 s		

## Data Link

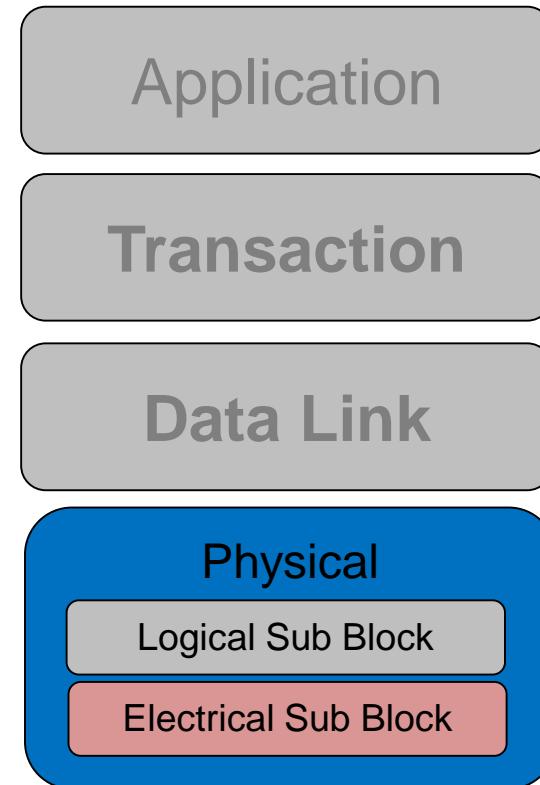
Packet	R→	5.0	DLLP	PM_Active_State_Request_L1	CRC 16	Time Delta	Time Stamp	
22		x1		0xEBO5	520.000 ns	0002 . 514 043 358 s		
Packet	R→	5.0	DLLP	UpdateFC-P	VC ID HdrFC DataFC	CRC 16	Idle	Time Stamp
23		x1		0 1 1024	0xC7CD	16.000 ns	0002 . 514 043 879 s	
Packet	R→	5.0	DLLP	UpdateFC-NP	VC ID HdrFC DataFC	CRC 16	Idle	Time Stamp
24		x1		0 1 1	0x6F0F	16.000 ns	0002 . 514 043 910 s	
Packet	R→	5.0	DLLP	UpdateFC-Cpl	VC ID HdrFC DataFC	CRC 16	Idle	Time Stamp
25		x1		0 1 1024	0x1102	48.000 ns	0002 . 514 043 942 s	
Packet	R→	5.0	DLLP	PM_Request_Ack	CRC 16	Time Delta	Time Stamp	
26		x1		0x930C	752.000 ns	0002 . 514 044 006 s		

## Physical

Logical Sub Block  
Electrical Sub Block

16 TS Packets	R←	8.0	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
3578669-3578699		x4		0 *	96	0 0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	0 0 7 0	2	18	5	4A ...	260.000 ns	0005 . 265 457 828 s
52 TS Packets	R←	8.0	TS2	Link Lane N_FTS	Training Control	Data Rate	Eq Control	TS2 Symbols		Time Delta	Time Stamp			
3578701-3578799		x4		0 *	96	0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	0 0 0	45 ...	458.750 ns	0005 . 265 458 088 s			

# Link Training Problems



# Questions to Ask When Debugging

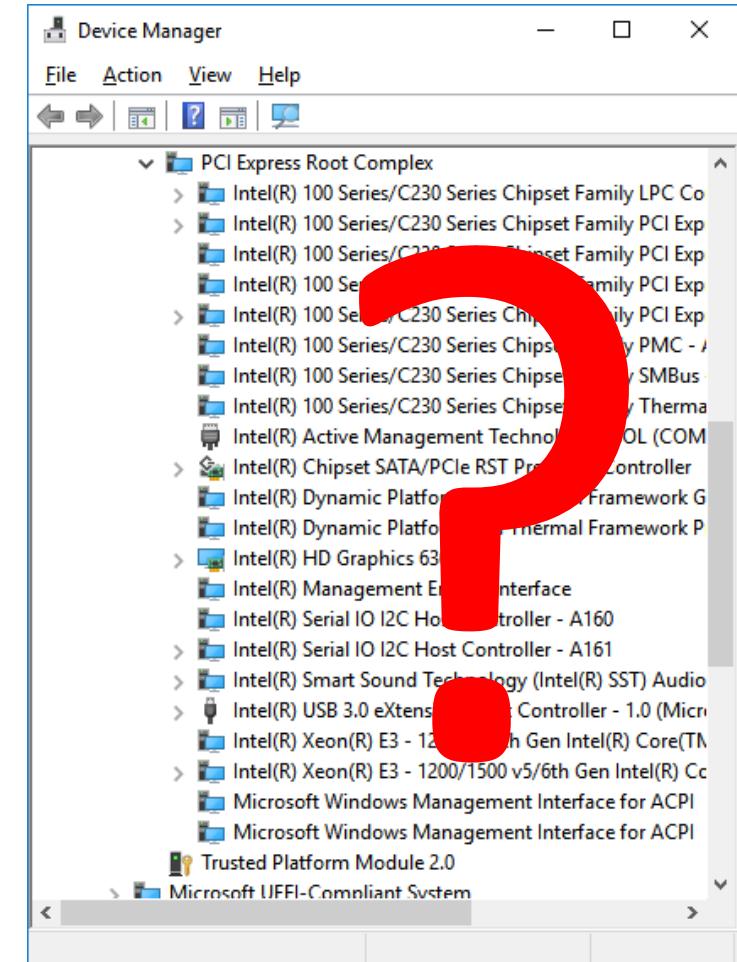


- 1. Does the PCIe link establish?**
- 2. Does the link try to establish then fail?**
- 3. Is the problem specific to a particular add in card/system slot combination?**

# My Device Does Not Show in Device Manager

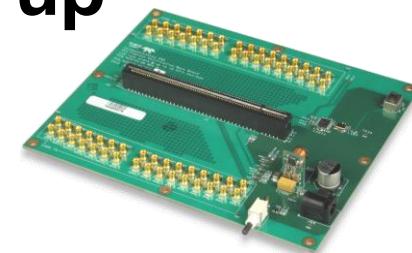
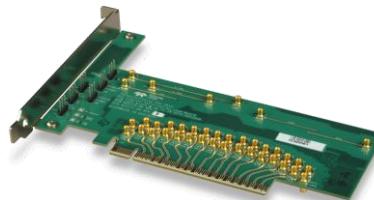


- **Is the Link Active?**
- **Do a snapshot trigger**
  - If the link is active, you will see Update\_FC packets and SKP ordered sets
  - If the link is not active, you may see TS1 or TS2 ordered sets, or Compliance Pattern
  - If the link is unstable, you may see Update\_FC and SKP ordered sets, with occasional TS1/TS2 sequences as the link goes to Recovery
- **Bottom Line: If the link is not active (in the L0 State), there is no way the device manager will see the device.**



# Signal Integrity

- **Is the Signal Integrity of the Link good on both sides?**
- **Signal Integrity problems can show in a whole manner of ways, from unstable links to devices just not showing up**



Question: Was Signal Integrity Testing only performed on Lane0?



# Snapshot or TS2 Trigger – Look for Clues

TS1 or TS2 Ordered Sets indicate that the link is not in L0

Packet	R←	8.0	x1	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp
6886787					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 7 0	7	45	11	4A ...	0.750 ns	0003 . 456 372 024 7 s
6886788					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 0 7 0	4	28	8	4A ...	15.500 ns	0003 . 456 372 025 5 s
6886789					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 041 0 s
6886790					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 0 7 0	4	28	8	4A ...	15.500 ns	0003 . 456 372 041 0 s
6886791					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 057 2 s
6886792					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 057 2 s
6886793					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 057 2 s
6886794					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 057 2 s
6886795					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 057 2 s
6886796					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 057 2 s
6886797					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 057 2 s
6886798					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 0 7 0	4	28	8	4A ...	15.500 ns	0003 . 456 372 090 5 s
6886799					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 106 7 s
6886800					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 122 2 s
6886801					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0 0					Stamp	0003 . 456 372 122 2 s

# PCIe Link Training from the Specification



PCI EXPRESS BASE SPECIFICATION, REV. 2.1

LTSSM State	Link Width	Link Speed	LinkUp	Link Training	Receiver Error	In-Band Presence <sup>44</sup>
Loopback	No action	Link speed may change on entry to Loopback from Configuration.	0b	0b	No action	1b
Hot Reset	No action	No action	0b	0b	Optional: Set on 8b/10b Error	1b

The state machine rules for configuring and operating a PCI Express Link are defined in the following sections.

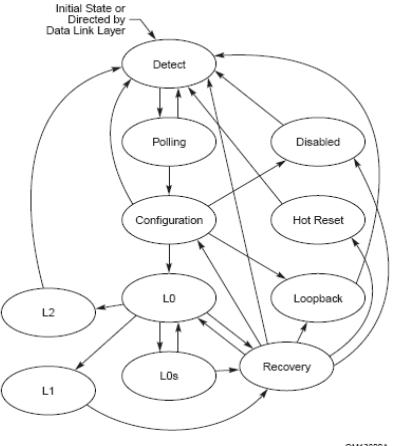


Figure 4-11: Main State Diagram for Link Training and Status State Machine

OM13600A

PCI EXPRESS BASE SPECIFICATION, REV. 2.1

#### 4.2.6.1. Detect

The Detect substate machine is shown in Figure 4-12.

##### 4.2.6.1.1. Detect.Quiet

- Transmitter is in an Electrical Idle state.
  - Note: The DC common mode voltage is not required to be within specification.
- 2.5 GT/s data rate is selected as the frequency of operation. If the frequency of operation was not 2.5 GT/s data rate on entry to this substate, the LTSSM must stay in this substate for at least 1 ms, during which the frequency of operation must be changed to the 2.5 GT/s data rate.
  - Note: This does not affect the advertised data rate in the TS1 and TS2 Ordered Sets.
- LinkUp = 0b (status is cleared).
- The directed\_speed\_change variable is reset to 0b. The upconfigure\_capable variable is reset to 0b. The idle\_to\_clock\_transitions variable is reset to 0b. The select\_deemphasis variable must be set to either 0b or 1b based on platform specific needs for the Downstream component and identical to the Selectable De-emphasis bit in the Link Control 2 register for an Upstream component.
  - Note that since these variables are defined with the 2.0 specification, pre-2.0 devices would not implement these variables and will always take the path as if the directed\_speed\_change and upconfigure\_capable variables are constantly reset to 0b and the idle\_to\_clock\_transitions variable is constantly set to 1b.
- The next state is Detect.Active after a 12 ms timeout or if Electrical Idle is broken on any Lane.

##### 4.2.6.1.2. Detect.Active

- The Transmitter performs a Receiver Detection sequence on all un-configured Lanes that can form one or more Links (see Section 4.3.1.8 for more information).
- Next state is Polling if a Receiver is detected on all unconfigured Lanes.
- Next state is Detect.Quiet if a Receiver is not detected on any Lanes.

# LTSSM State Machine

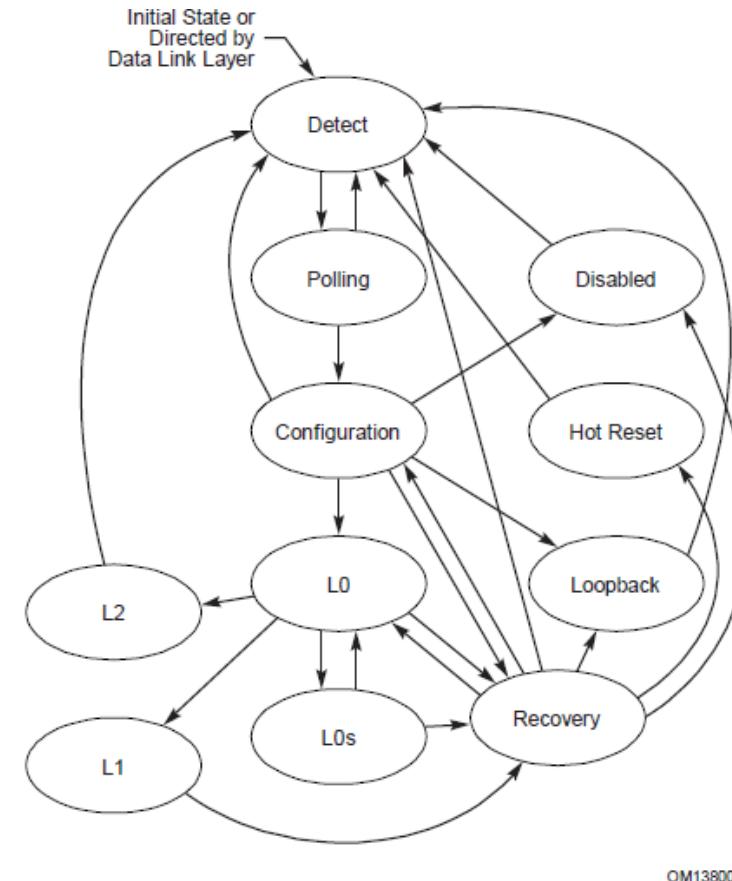
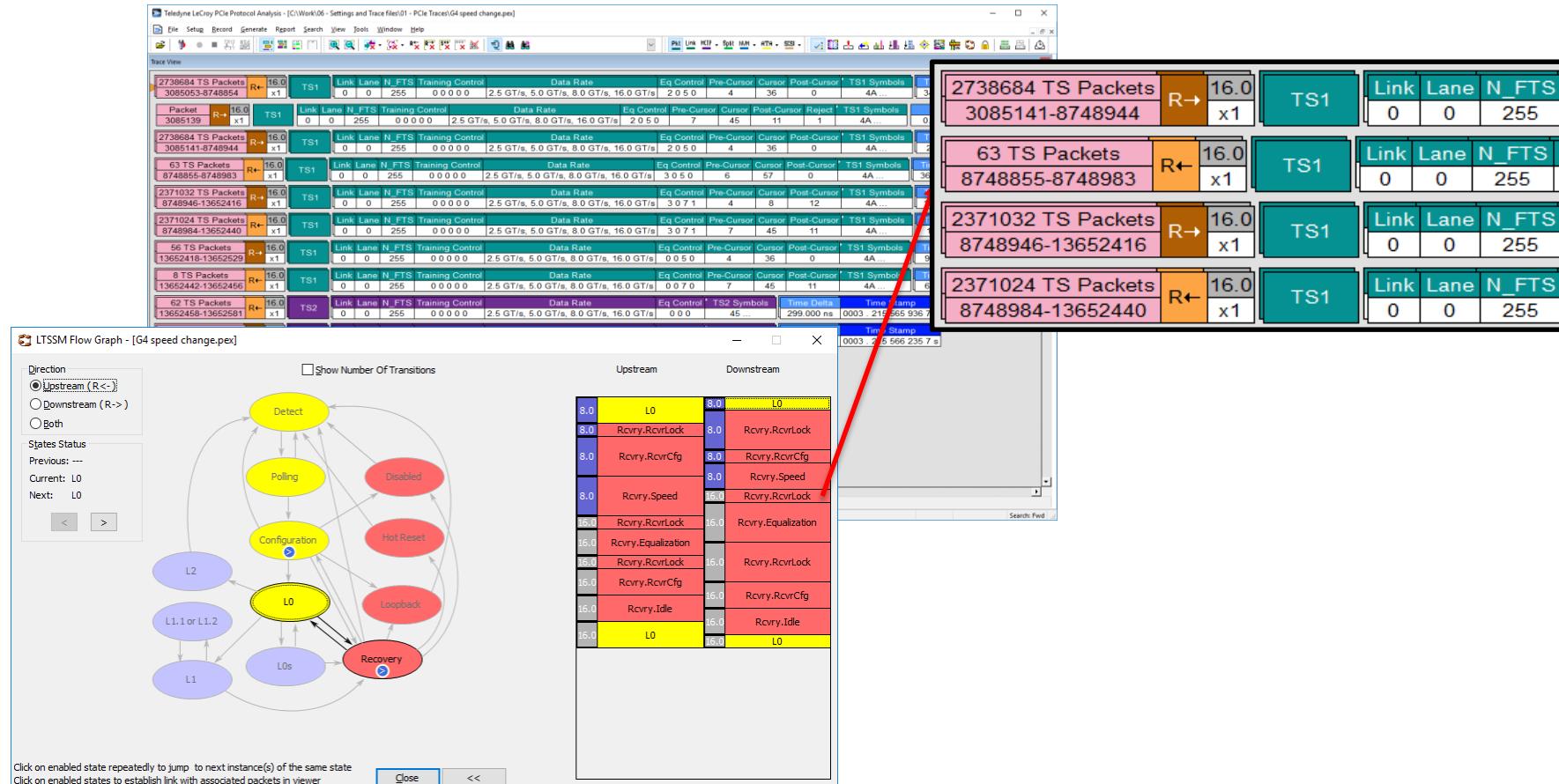


Figure 4-23: Main State Diagram for Link Training and Status State Machine

# PCIe Link Training Debug

- LTSSM state changes appear as 1000's of ordered sets



# Example: Link is Not Running at Maximum Speed

Trace View											
Packet	R→	8.0		SKIP	SKIP Symbols	END	P	LFSR Symbols	Time Delta	Time Stamp	
13340		x1			AAAAAAA... A	E1	0	56 45 DE	1.266 us	0003 . 159 615 586 5 s	
13342		x1		SKIP	SKIP Symbols	END	P	LFSR Symbols	Idle	Time Stamp	
13343		x1	DLLP		UpdateFC-P	VC ID	HdrScale	HdrFC	DataScale	DataFC	CRC 16 0xC7CD
13344		x1	DLLP		UpdateFC-NP	VC ID	HdrScale	HdrFC	DataScale	DataFC	CRC 16 0x6F0F
13345		x1	DLLP		UpdateFC-Cpl	VC ID	HdrScale	HdrFC	DataScale	DataFC	CRC 16 0x1102
13347		x1	SKIP		SKIP Symbols	END	P	LFSR Symbols	Time Delta	Time Stamp	
					AAAAAAA... A	E1	1	34 AB D2	1.266 us	0003 . 159 621 615 2	
13349		x1	SKIP		SKIP Symbols	END	P	LFSR Symbols	Time Delta	Time Stamp	
					AAAAAAA... A	E1	1	61 28 31	3.987 us	0003 . 159 622 881	
13351		x1	EIEOS		EIEOS Symbols				Idle	Time Stamp	
					00 FF				0.250 ns	0003 . 159 626 977 7 s	
97 TS Packets	R→	8.0	x1	TS1	Link	Lane	N_FTS	Training Control	Data Rate		Time Delta 93.000 ns
13352-13546					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s		
Packet	R←	8.0	x1	EIEOS	EIEOS Symbols				Idle	Time Stamp	
13359					00 FF				0.250 ns	0003 . 159 626 977 7 s	
7 TS Packets	R→	8.0	x1	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Time Delta 113.750 ns
13361-13373					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 7 0	0003 . 159 626
9 TS Packets	R→	8.0	x1	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Time Delta 146.250 ns
13375-13391					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 0 7 0	0003 . 159 626
208 TS Packets	R→	8.0	x1	TS2	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Time Delta 1.262 us
13393-13820					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 0 0	0003 . 159 627 254 0 s
130 TS Packets	R→	8.0	x1	TS2	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Time Delta 2.194 us
13548-13816					0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 1 0 5	0003 . 159 628 515 7 s
Packet	R→	8.0	x1	EIOS	EIOS Symbols				Time Stamp		
13818					66 66 66 ...				0003 . 159 630 709 5 s		

Do both sides advertise higher speeds— Was a speed change initiated?

# Signal Aggressor Counter Actions



# Dynamic Equalization

- The PCI Express 4.0 Specification allows for 8GT/s and 16GT/s data rates
- In order to establish a link reliably at 8GT/s and above, the protocol allows for dynamic equalization of the link during the speed change in the Recovery state
- This is a 4 phase process and happens in the Recovery.Equalization substate

# Recovery Substate Machine

Recovery.Equalization  
Substate

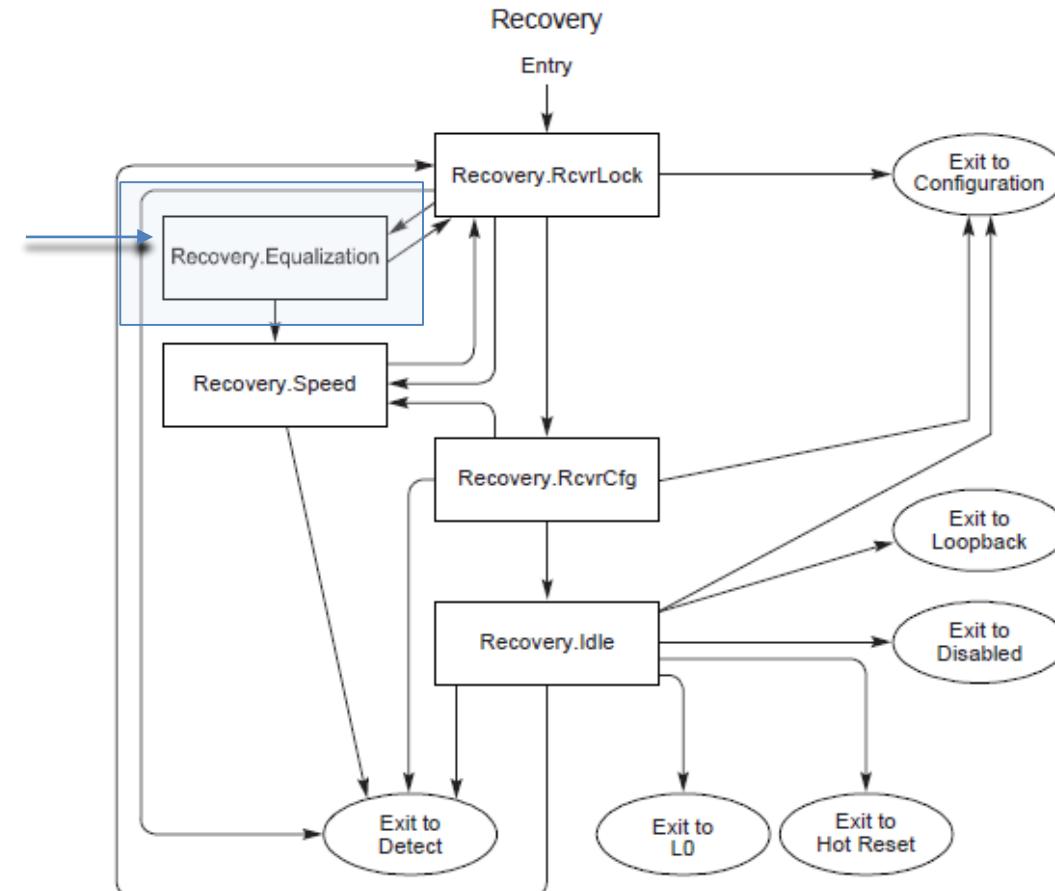


Figure 4-26: Recovery Substate Machine

A-0522A

# Equalization Process

- **See Section 4.2.3 of PCIe 4.0 Base Specification**
- **The Link equalization procedure enables components to:**
  - Adjust the Transmitter and the Receiver setup of each Lane to improve the signal quality
  - Meet the requirements specified in the physical layer specification, when operating at 8GT/s and higher data rates.
- **The procedure must be executed during the first data rate change to 8GT/s as well as the first change to all data rates greater than 8GT/s.**

# Equalization Process Part 1 – 2.5GT/s – 8GT/s

- Process to change to 16GT/s starts with a speed change to 8GT/s, without advertising 16GT/s
- “The procedure must be executed during the first data rate change to 8GT/s as well as the first change to all data rates greater than 8GT/s.”

94 TS Packets 22080-22267	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Fs LF	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp		
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	1 0 7 0	24 7	5	4A ...	1.099 us	0005 . 537 780 140 s		
62 TS Packets 22218-22344	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Fs LF	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp		
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	1 0 7 0	24 7	5	4A ...	408.500 ns	0005 . 537 781 240 s		
63 TS Packets 22269-22397	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	2 0 7 0	2	17	5	4A ...	631.500 ns	0005 . 537 781 648 s	
1180523 TS Packets 22346-2463752	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	2 0 7 1	11	5	8	4A ...	434.750 ns	0005 . 537 782 280 s	
1180522 TS Packets 22399-2463806	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	Reject	TS1 Symbols	Time Delta	Time Stamp
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	2 0 7 0	0	0	0	1	4A ...	19.836 ms	0005 . 537 782 714 s
62 TS Packets 2463754-2463879	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	3 0 7 0	2	17	5	4A ...	448.750 ns	0005 . 557 618 726 s	
1180528 TS Packets 2463808-4905228	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	3 0 7 1	4	8	12	4A ...	591.250 ns	0005 . 557 619 174 s	
1180530 TS Packets 2463881-4905302	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	Reject	TS1 Symbols	Time Delta	Time Stamp
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	3 0 7 0	0	0	0	1	4A ...	19.836 ms	0005 . 557 619 766 s
64 TS Packets 4905230-4905360	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 0 7 0	2	17	5	4A ...	611.500 ns	0005 . 577 455 728 s	
10 TS Packets 4905304-4905325	R→ 8.0 x1	TS1	Link Lane N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 0 7 0	2	17	5	4A ...	178.750 ns	0005 . 577 456 340 s	
70 TS Packets 4905327-4905440	R→ 8.0 x1	TS2	Link Lane N_FTS	Training Control	Data Rate	Eq Control	TS2 Symbols			Time Delta	Time Stamp		
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 1 0	45 ...			286.250 ns	0005 . 577 456 518 s		
25 TS Packets 4905362-4905409	R→ 8.0 x1	TS2	Link Lane N_FTS	Training Control	Data Rate	Eq Control	TS2 Symbols			Time Delta	Time Stamp		
			0 0 255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s	0 1 0	45 ...			390.250 ns	0005 . 577 456 806 s		

# Equalization Process Part 2 – 8GT/s – 16GT/s



- **Immediately following the transition from Recovery to L0, after the initial data rate change to 8GT/s, the Downstream Port is required to:**
  - transition from L0 to Recovery
  - advertise 16GT/s data rate support
  - change the data rate to 16GT/s and perform the 16GT/s equalization procedure
- **If**
  - both components advertised that they are capable of 16GT/s during the initial Link negotiation,
  - neither component detected problems with its 8GT/s equalization settings and it intends to perform a 16GT/s equalization procedure using the autonomous mechanism.

# Equalization Process Part 2 – 8GT/s – 16GT/s

- Speed Change is initiated to change to speed above 8GT/s

TS1 advertises 16GT/s Data Rate



97 TS Packets	R→	8.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp
4905444-4905606	R←	x1		0	0	255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 0 7 0	2	17	5	4A ...	639.750 ns	0005.577 458 202 s
5 TS Packets	R→	8.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp
4905486-4905494	R←	x1		0	0	255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 7 0	2	17	5	4A ...	81.250 ns	0005.577 458 842 s
92 TS Packets	R→	8.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp
4905496-4905683	R←	x1		0	0	255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 0 7 0	2	17	5	4A ...	920.000 ns	0005.577 458 924 s
191 TS Packets	R→	8.0	TS2	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	TS2 Symbols			Time Delta	Time Stamp	
4905608-4905980	R←	x1		0	0	255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 1 0 0 7	45 ...			623.750 ns	0005.577 459 844 s	
130 TS Packets	R→	8.0	TS2	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	TS2 Symbols			Time Delta	Time Stamp	
4905685-4905954	R←	x1		0	0	255	0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s, Speed Change	0 1 0 0 7	45 ...			2.194 us	0005.577 460 468 s	

# Equalization Process Part 2 – 8GT/s – 16GT/s

**Equalization Phase 1**

**Equalization Phase 2**

**Equalization Phase 3**

**Link Reaches L0 at 16GT/s**

826884 TS Packets		R→	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	FS	LF	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
6887387-8481182		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	1 0 5 0	40	12	0	4A ...	944.604 us	0003 . 456 433 748 0 s	
714408 TS Packets		R←	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
7003649-8481095		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 5 0	6	57	0	4A ...	6.002 ms	0003 . 457 378 352 0 s	
714440 TS Packets		R←	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	FS	LF	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
8481097-9958609		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	1 0 5 0	63	21	0	4A ...	356.000 ns	0003 . 463 380 488 5 s	
3453124 TS Packets		R→	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
8481184-15622500		x1							2.5 GT/s, 8.0 GT/s, 16.0 GT/s	2 0 5 0	4	36	0	4A ...	6.002 ms	0003 . 463 380 844 5 s	
2738685 TS Packets		R←	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
9958611-15622413		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	2 0 5 0	4	36	0	4A ...	23.009 ms	0003 . 469 382 895 0 s	
63 TS Packets		R→	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
15622415-15622542		x1							2.5 GT/s, 8.0 GT/s, 16.0 GT/s	3 0 5 0	6	57	0	4A ...	356.250 ns	0003 . 492 392 116 7 s	
2371030 TS Packets		R→	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
15622502-20525967		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	3 0 7 1	4	8	12	4A ...	172.000 ns	0003 . 492 392 473 0 s	
2371021 TS Packets		R←	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
15622544-20525991		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	3 0 7 1	7	45	11	4A ...	19.920 ms	0003 . 492 392 645 0 s	
55 TS Packets		R→	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
20525969-20526080		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 5 0	4	36	0	4A ...	103.000 ns	0003 . 512 312 821 5 s	
8 TS Packets		R←	16.0	TS1	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	Pre-Cursor	Cursor	Post-Cursor	TS1 Symbols	Time Delta	Time Stamp	
20525993-20526008		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 7 0	7	45	11	4A ...	65.000 ns	0003 . 512 312 924 5 s	
62 TS Packets		R→	16.0	TS2	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	TS2 Symbols	Time Delta	Time Stamp				
20526010-20526133		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 0 0	45	...		297.000 ns	0003 . 512 312 989 5 s		
26 TS Packets		R→	16.0	TS2	Link	Lane	N_FTS	Training Control	Data Rate	Eq Control	TS2 Symbols	Time Delta	Time Stamp				
20526082-20526132		x1			0	0	255	0 0 0 0 0	2.5 GT/s, 5.0 GT/s, 8.0 GT/s, 16.0 GT/s	0 1 0	45	...		209.250 ns	0003 . 512 313 286 5 s		
Packet		R→	16.0	SKIP	SKIP Symbols	END	P	FR	SR	CRC	MP	UM	MT	RN	Payload	Time Delta	Time Stamp
20526134		x1			AAAAAAA... AAAAAAAA...	78	0	0	0	0	0	0	3	1	56	5.750 ns	0003 . 512 313 495 7 s
Packet		R←	16.0	SKIP	SKIP Symbols	END	P	FR	SR	CRC	MP	UM	MT	RN	Payload	Time Delta	Time Stamp
20526135		x1			AAAAAAA... AAAAAAAA...	78	0	0	0	0	0	0	7	0	156	0.500 ns	0003 . 512 313 501 5 s
Packet		R→	16.0	SDS	SDS Symbols	Time Delta	Time Stamp										
20526136		x1			E1 55 55 55 55 ...	7.500 ns	0003 . 512 313 502 0 s										
Packet		R←	16.0	SDS	SDS Symbols	Time Delta	Time Stamp										
20526137		x1			E1 55 55 55 55 ...	8.750 ns	0003 . 512 313 509 5 s										

- Normally the Link should reach the L0 state and remain there
- If the Link frequently enters the Recovery state, then this may be a symptom of various different conditions

136 LTSSM states were logged:				
L0	x8	16.0 GT/s	2.420 ms	
-----				
Recovery.RcvrLock	x8	16.0 GT/s	1.504 us	
Recovery.RcvrCfg	x8	16.0 GT/s	184 ns	
Recovery.Idle	x8	16.0 GT/s	544 ns	
L0	x8	16.0 GT/s	24.897 ms	
-----				
Recovery.RcvrLock	x8	16.0 GT/s	848 ns	
Recovery.RcvrCfg	x8	16.0 GT/s	224 ns	
Recovery.Idle	x8	16.0 GT/s	496 ns	
L0	x8	16.0 GT/s	23.004 ms	
-----				
Recovery.RcvrLock	x8	16.0 GT/s	840 ns	
Recovery.RcvrCfg	x8	16.0 GT/s	232 ns	
Recovery.Idle	x8	16.0 GT/s	488 ns	
L0	x8	16.0 GT/s	9.655 ms	
-----				
Recovery.RcvrLock	x8	16.0 GT/s	856 ns	
Recovery.RcvrCfg	x8	16.0 GT/s	192 ns	
Recovery.Idle	x8	16.0 GT/s	536 ns	
L0	x8	16.0 GT/s	146.886 ms	

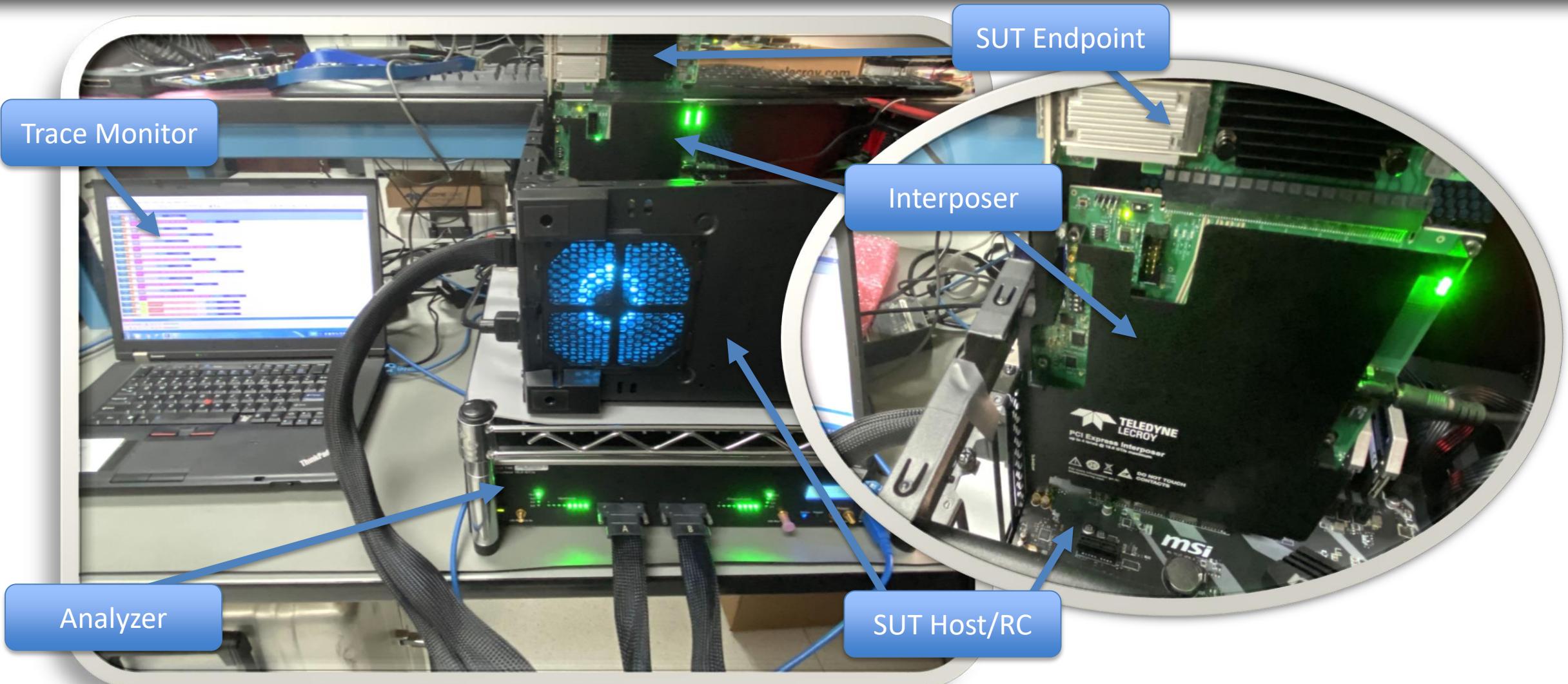
Look at the period  
between entries to  
Recovery

# Potential Equalization Problems

- Link does not get to 8GT/s , 16GT/s or 32GT/s
- Link gets to 8GT/s or 16GT/s but is unstable
- Equalization fails because of wrong coefficients
- Does the equalization process complete?
- Protocol Analyzer is the only way to get visibility

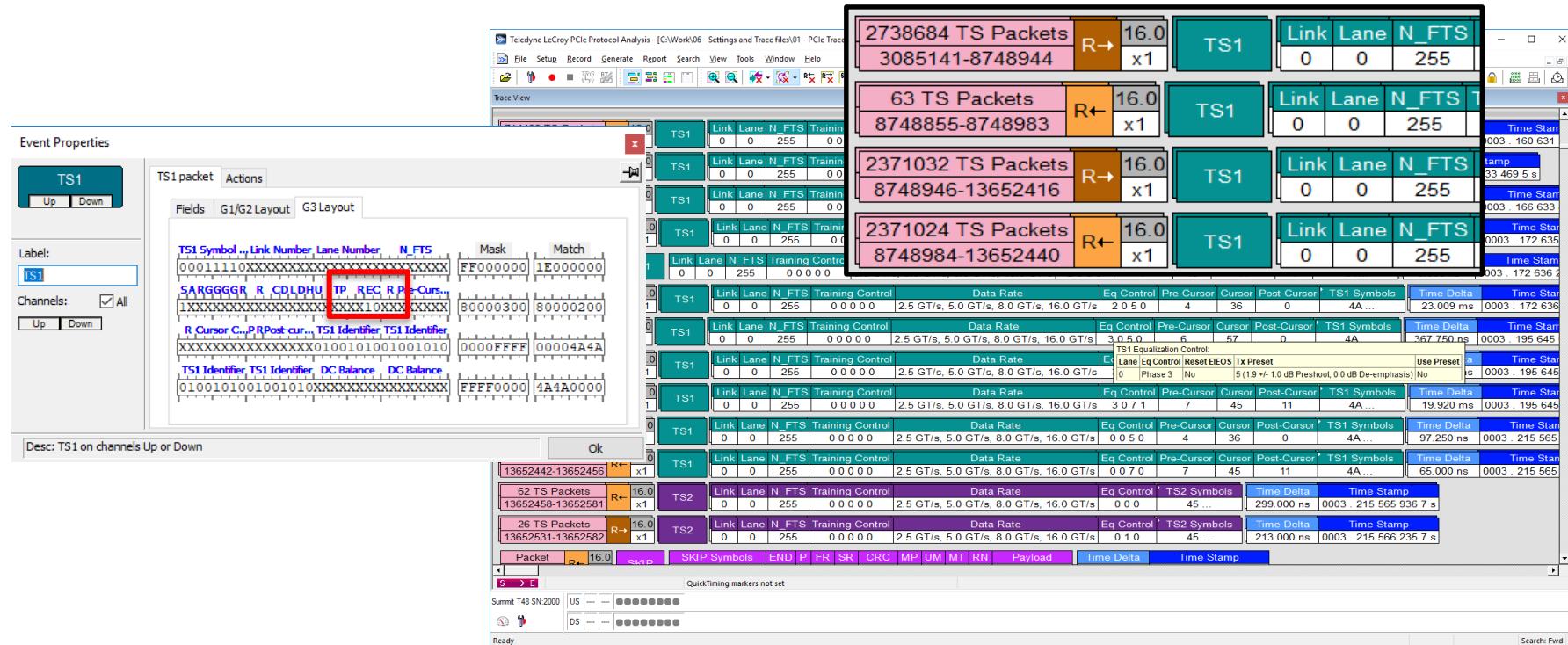


# Getting a Signal the Easy Way

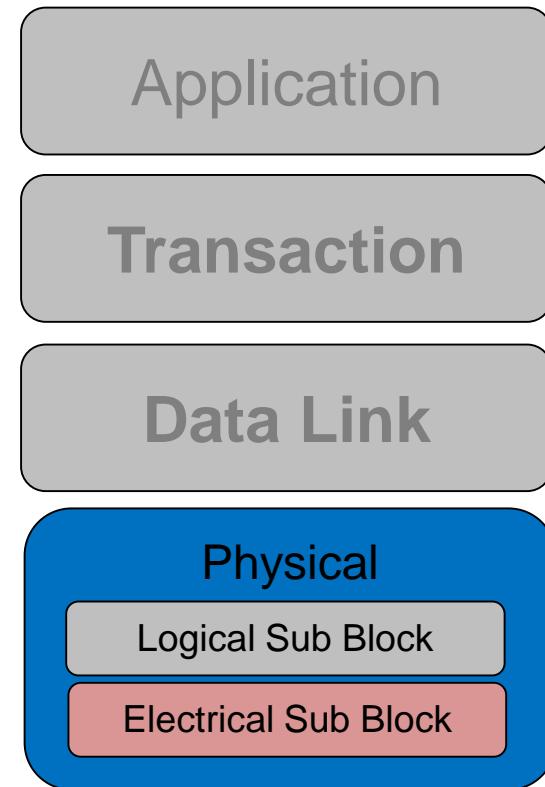


# How Do I Debug Equalization?

- Use the EC field as a trigger – this will help navigate the equalization
- Compress the Training Sequences helps digest the process – there are thousands of them!



# Polarity Issues



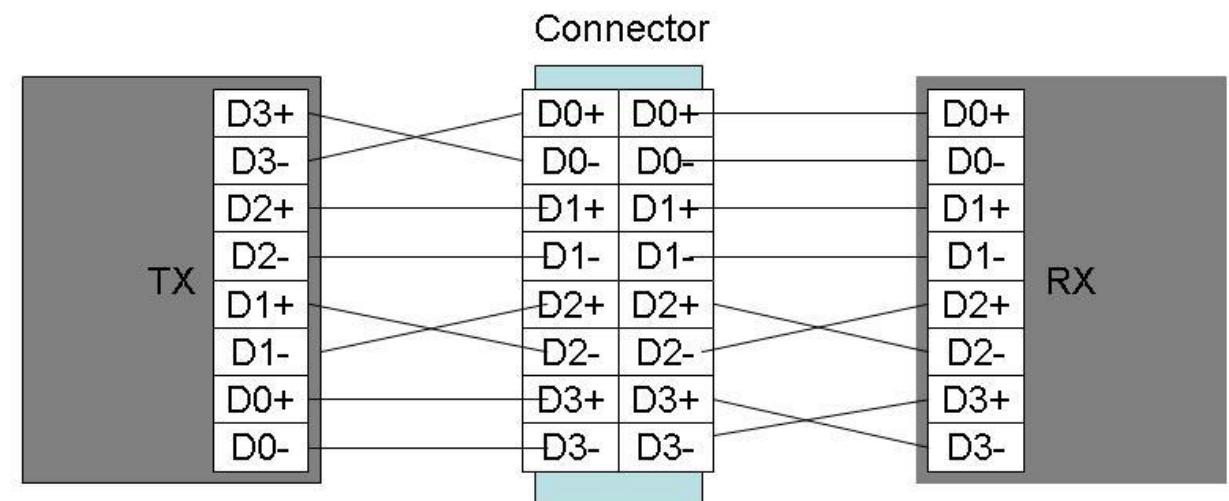
# Polarity Inversion

- PCI Express allows Polarity Inversion
- New form factors and adapters are exhibiting this more often
- This can often be challenging to work out when setting up a Protocol Analyzer
- It is quite common to see different polarity settings on different slots in the same system

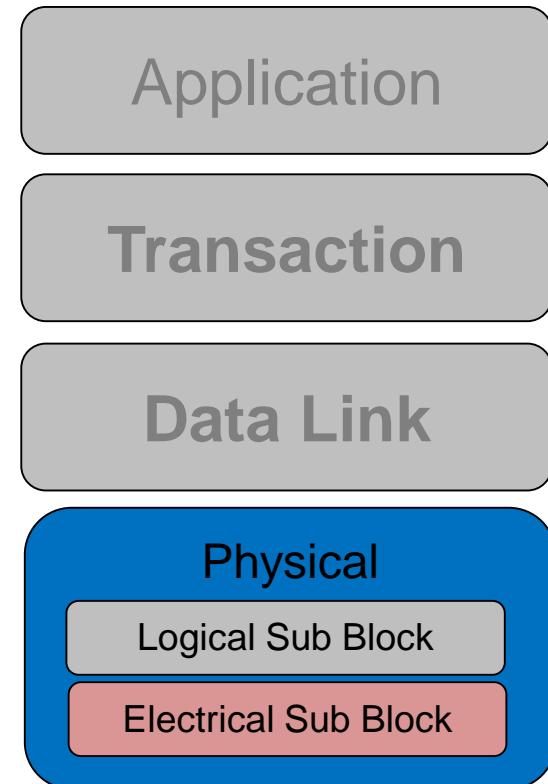
# Polarity Issues

- **Normally Polarity can be auto detected**
  - It is fixed for any particular slot or card
  - If Power management (L0s/L1) is enabled, it is always better to use manually assigned polarity settings to improve lock times exiting Electrical Idle
- **Incorrect Polarity settings will result in a trace full of errors**

Depending on the system layout, type of probe being used, it may see polarity different to the system under test

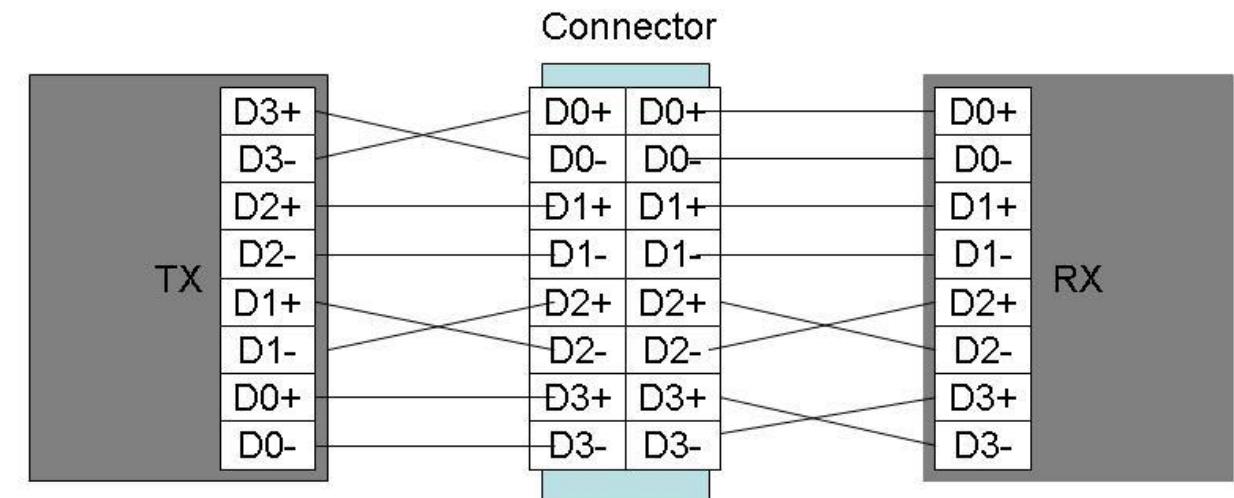


# Lane Reversal



# Lane Reversal

- Many implementations are now using Lane Reversal
- Especially true in Server platforms with U.2 connections
- This often causes analysis problems
- How to identify Lane Reversal?

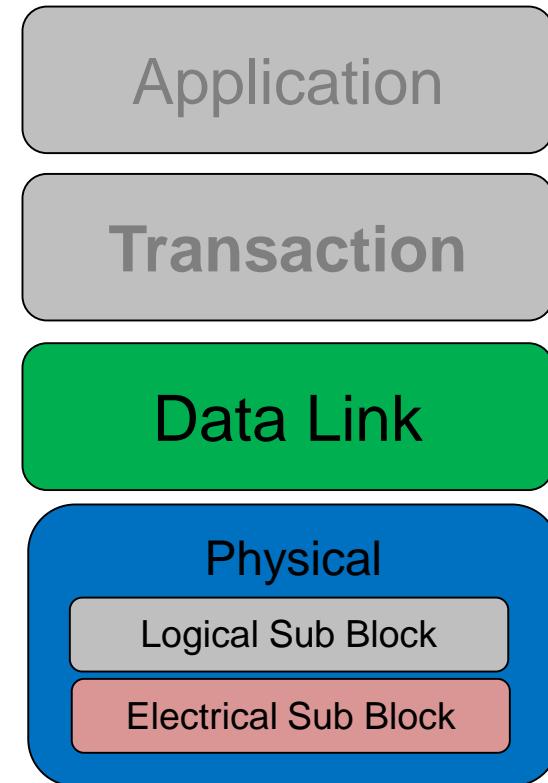


# Lane Reversal

- **The trick when looking for Lane Reversal is:**
  - Trigger on TS2 Ordered Set
  - Look in the Lane Number field
  - This should read 0,1,2,3
  - If it reads 3,2,1,0 then the link has Lane Reversal

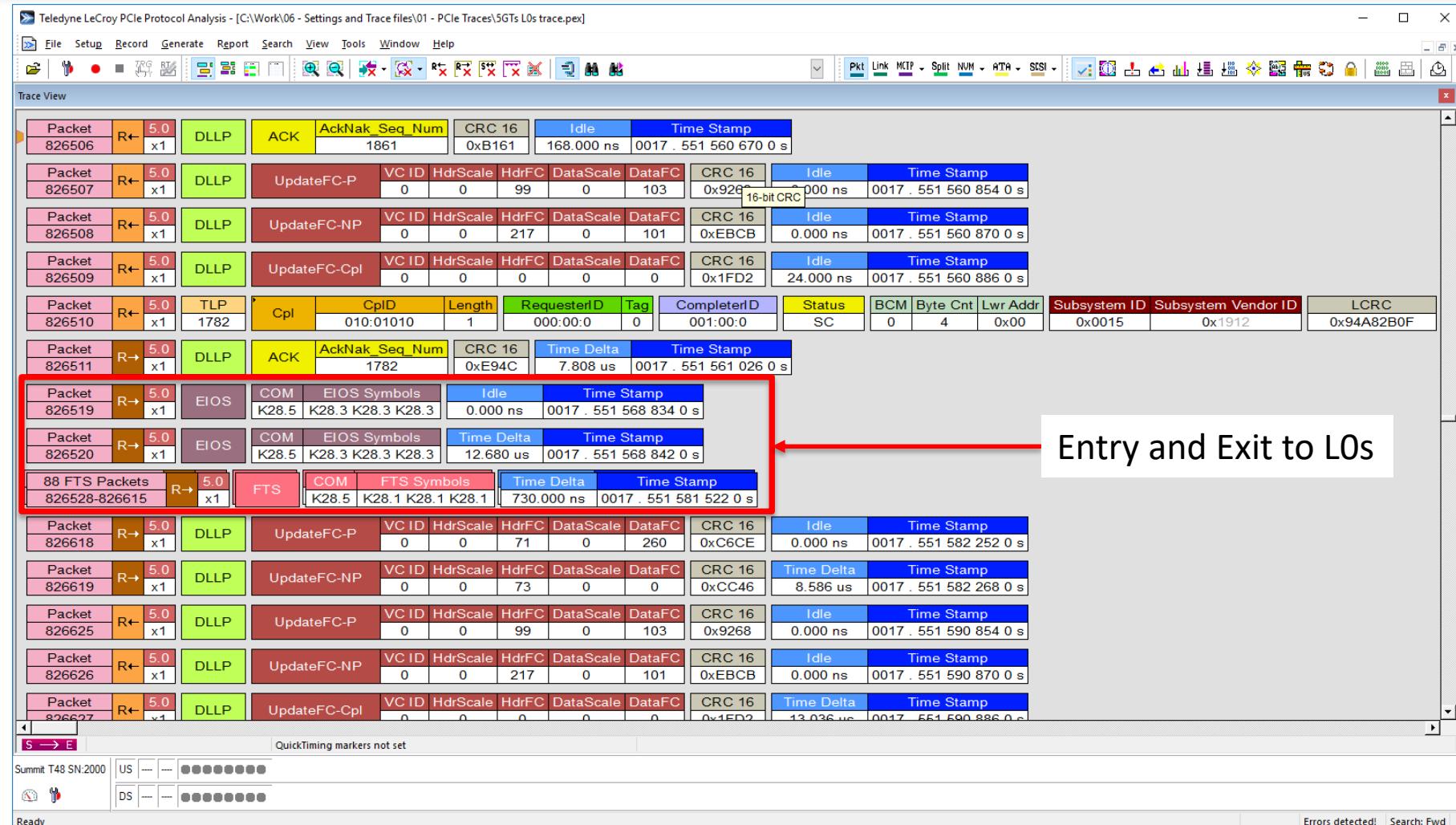
Packet	R←	2.5	x4	TS2	COM	Lin	Lane	I_FTS	Training Control	Data Rate	TS2 Symbols	Time Delta	Time Stamp
4437					K28.5	0	0	255	0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	D05.2	0.000 ns	0007 . 165 541 810 s
					K28.5	0	1	255	0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	D05.2		
					K28.5	0	2	255	0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	D05.2		
					K28.5	0	3	255	0 0 0 0	2.5 GT/s, 5 GT/s, 8 GT/s	D05.2		

# Power Management Problems



- **It can be challenging to debug L0s problems, it is important to set things up properly**
  - The exit latency from L0s is very short
  - Give the analyzer the best chance of locking by disabling auto detection such as speed, link width and polarity
  - Make sure the analyzer is set up properly with L0s turned off and ensure clean captures

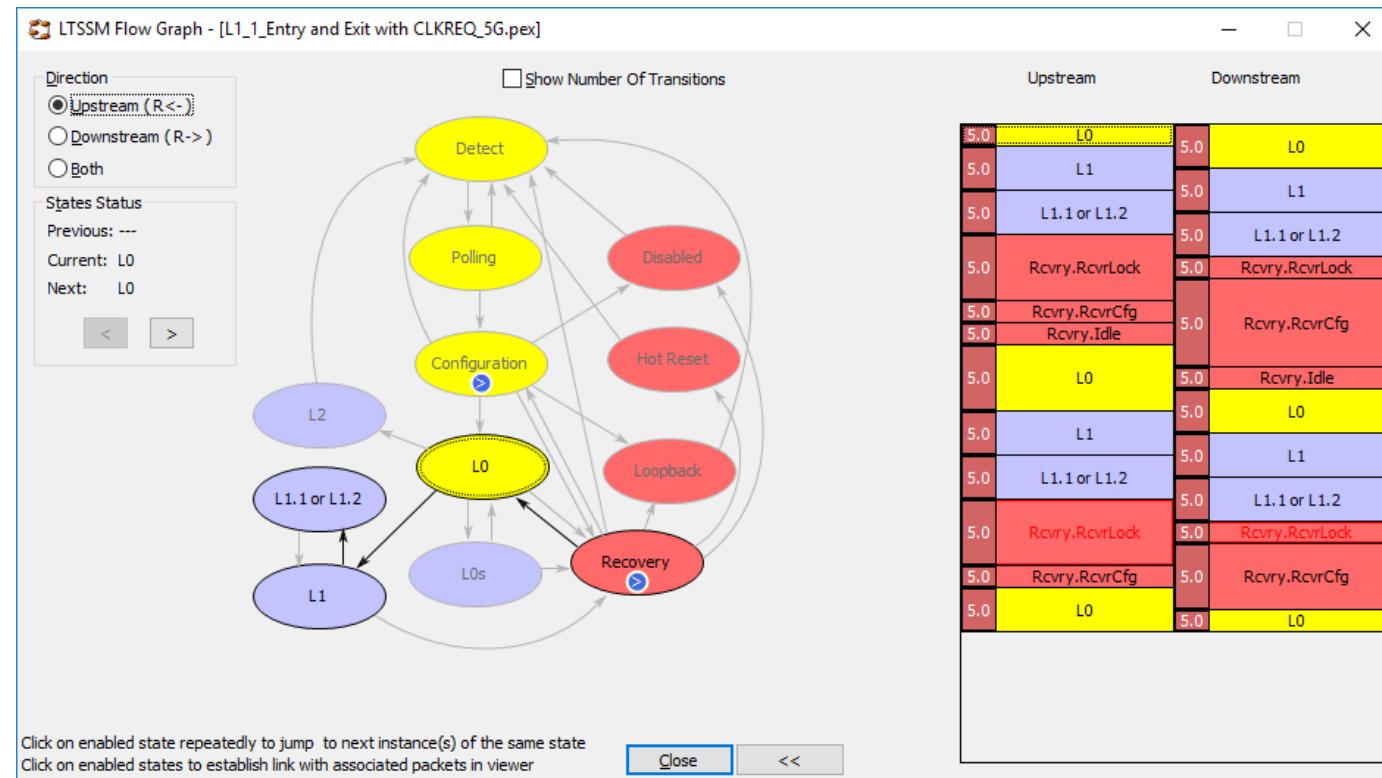
# ASPM – L0s



# L1 and L1 Substates

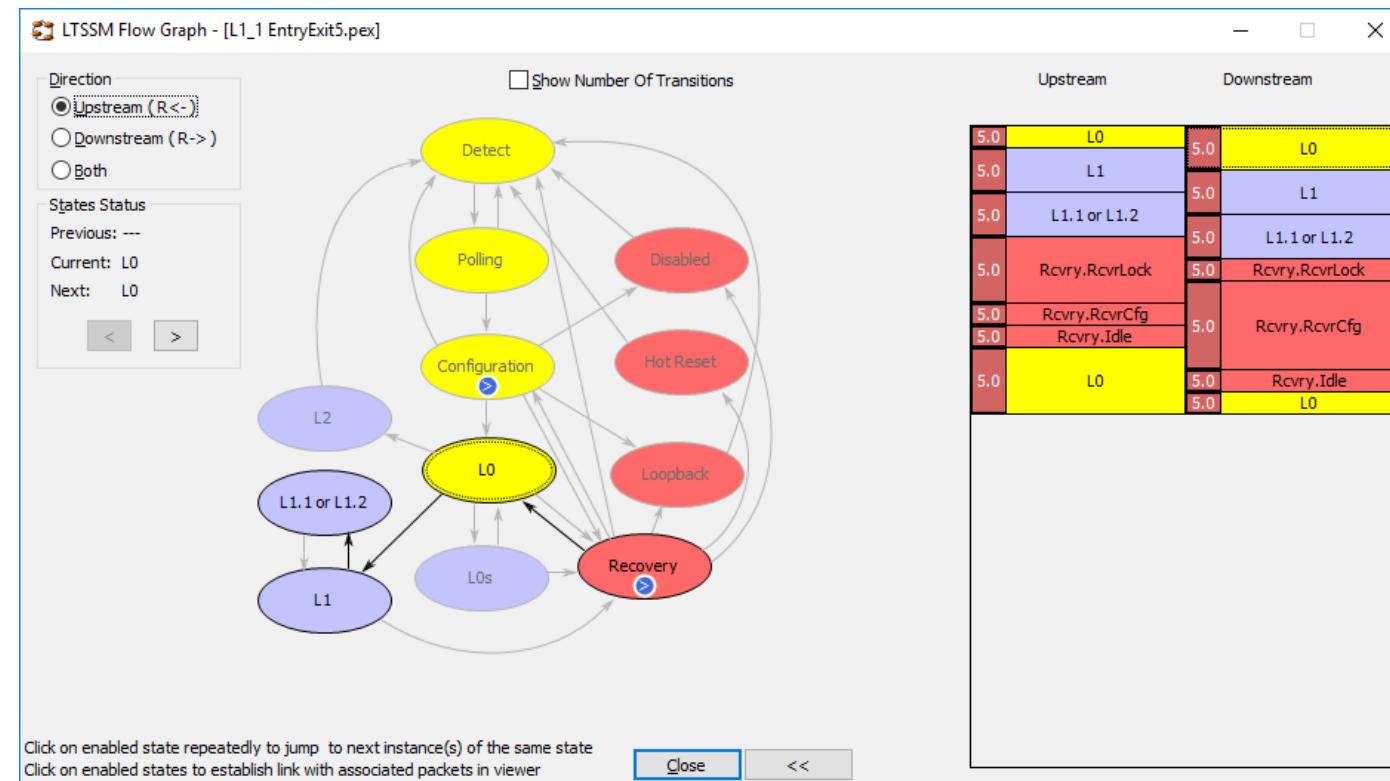
- **L1 is a more aggressive form of power management than L0s, and the L1 Substates ECN takes that further**
- **There are 2 mechanisms to enter L1 - ASPM L1 and PCIPM L1**
  - Each have a different entry mechanism
  - PCIPM uses a config write from the RC to the PMSCR register on the device to initiate the transition to L1
  - ASPM L1 the downstream component indicates the desire to enter the L1 state, and sends PM\_Active\_State\_Request\_L1 to the root, the root acknowledges with PM\_Request\_ACK DLLP

- Advanced post processing shows how many times each state was entered



# ASPM – L1.1 and L1.2

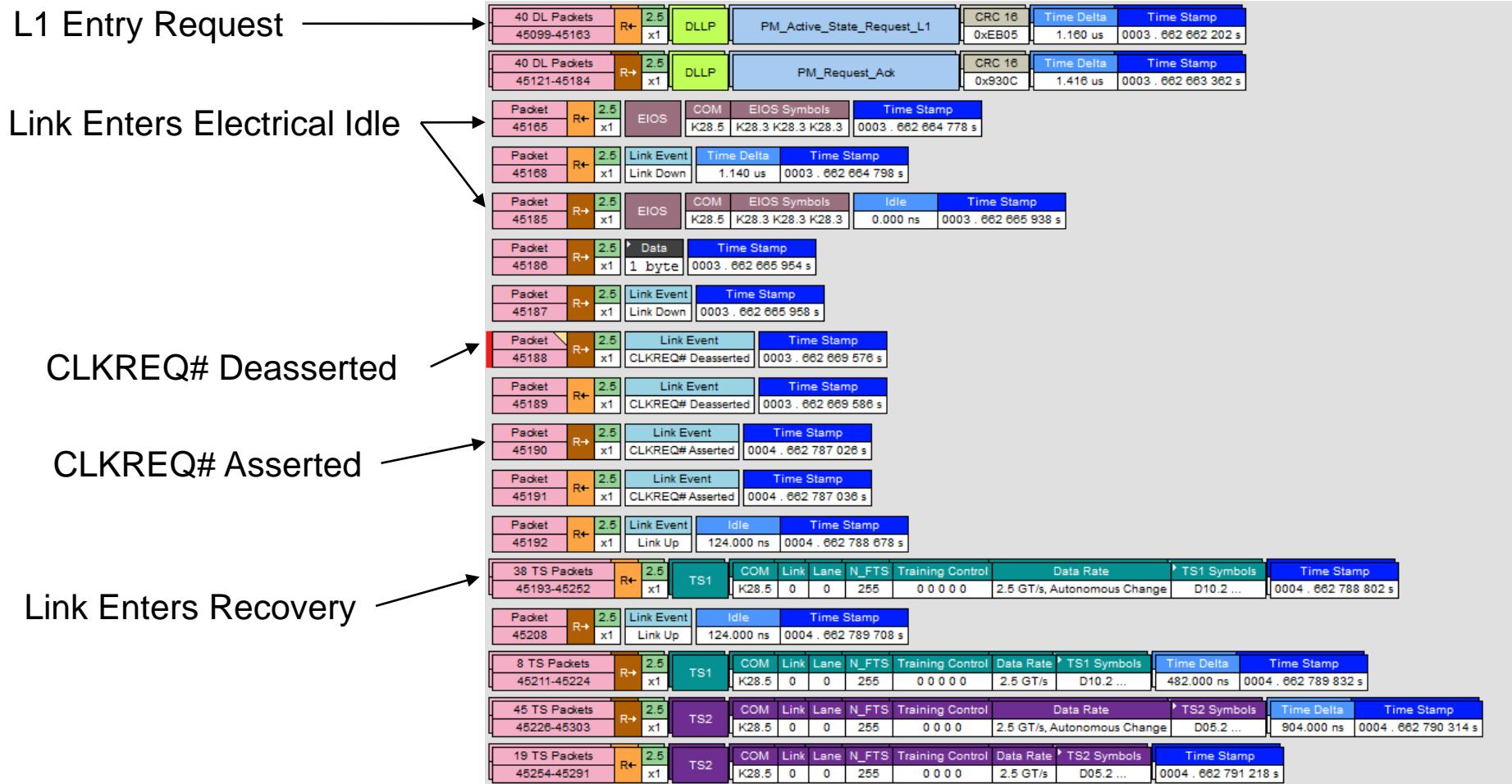
- L1.1 and L1.2 are substates of L1



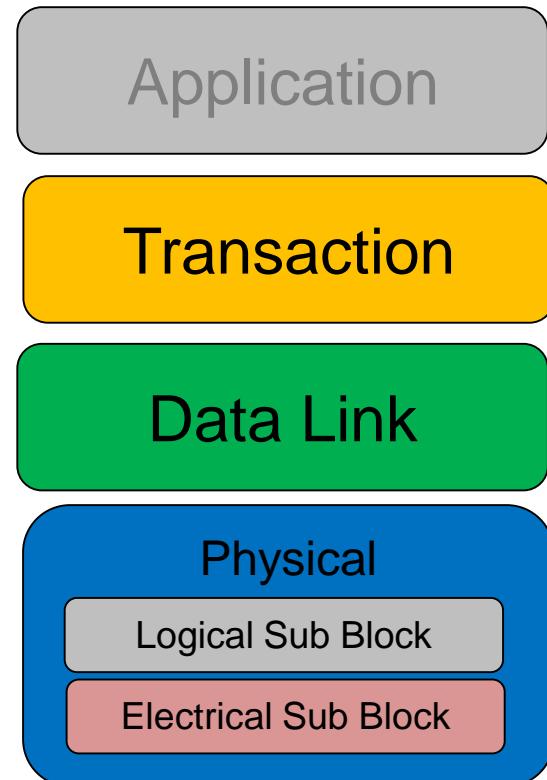
# L1 Substates

- A device indicates its support of L1 substates in the configuration space
- L1 PM Substates is considered enabled on a Port when any combination of the ASPM L1.1 Enable, ASPM L1.2 Enable, PCI-PM L1.1 Enable and PCI-PM L1.2 Enable bits associated with that Port are Set

# How Does it Look on a Protocol Analyzer?



# Reset Problems



# PCIe Reset Mechanisms



- **There are different types of reset mechanism**
  - Cold, Warm, Hot reset
  - Function Level reset

# Reset Mechanisms

- **Conventional Reset includes all reset mechanisms excluding Function Level Reset**
- **All form factors need a mechanism to return all Port state to the initial condition, called Fundamental Reset**
  - Must occur following application of power to the component – called a “Cold reset”
  - Can be done without removing power – “Warm reset”
  - In band mechanism for propagating Conventional reset across a Link – “Hot reset”
- **See section 6.6.1 in PCI Express Specification**

# PCIe Reset Mechanisms - FLR

- **Function Level Reset**
- **The FLR mechanism enables software to quiesce and reset Endpoint hardware with Function-level granularity**
- **FLR applies on a per Function basis. Only the targeted Function is affected by the FLR operation.**
- **The Link state must not be affected by an FLR**
- **Controls that enable the Function to initiate requests on PCI Express are cleared, including Bus Master Enable, MSI Enable etc.**
- **The Function must complete the FLR within 100 ms.**
- **FLR is optional in PCIe spec but required in NVMe**

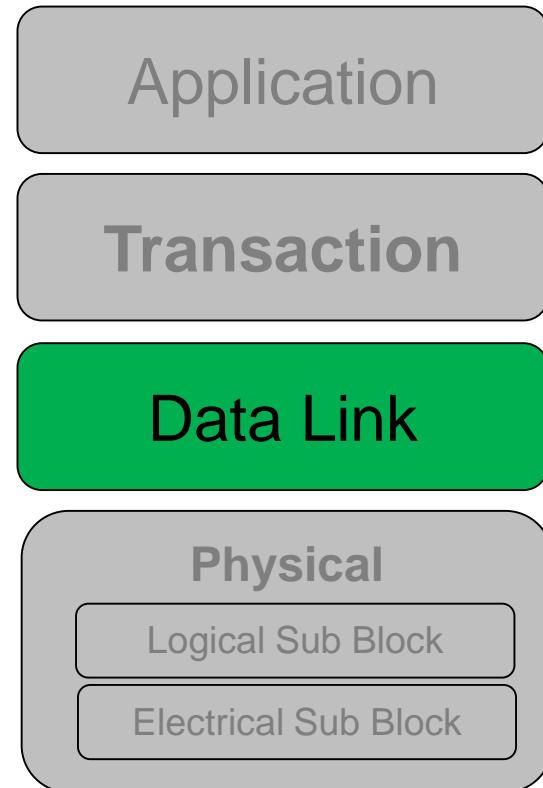
# Typical Problems with Reset



- **Possible problems that may occur**

- Registers not reset properly
- Errors reported during Reset
- Link fails after Reset

# Flow Control Problems



# Data Link Layer Issues

- **Some problems are related to flow control mechanisms**
  - Most DLL FC problems will not be discovered by the current compliance tests
  - Many DLL FC problems are not discovered until after a product is released
  - Problems may only show up with specific combination of Root and End Point that is not tested during compliance testing
  - Data Link Layer problems often result in performance issues

# Detecting Credit Errors

PCI SIG Developers Conference 2019

Copyright © 2019 PCI-SIG® - All Rights Reserved

**LOOK HERE!**

Frame	Type	VC ID	HdrFC	DataFC	CRC 16	Time Delta	Time Stamp
1404	InitFC2-Cpl	0	0	0	0xA2ED	0.000 ns	0025 . 257 597 748 s
1405	InitFC2-P	0	2	16	0xAE65	40.000 ns	0025 . 257 597 780 s
1406	UpdateFC-P	0	8	16	0x315A	FC-P	0 264 4112
1407	UpdateFC-NP	0	8	8	0xD3FA	FC-NP	0 264 4104
1415	UpdateFC-P	0	3	17	0x5DAF	FC-P	0 2 16
						FC-Cpl	0 264 4112
						FC-NP	0 264 4104

The screenshot shows a PCI-SIG trace analysis interface. The menu bar includes Report, Search, View, Tools, Window, Help, Toolbars, Analyzer Network Chat Bar, Status Bar, Real-time Statistics, Zoom In, Zoom Out, Wrap, FC Credits (which is selected), FC Credits Setup..., Compact View, Packet Level, Link Transaction Level, and Split Transaction Level. The main window displays a sequence of FC frames with their respective fields: VC ID, HdrFC, DataFC, and CRC 16. A yellow arrow points to the last frame (Packet 1415) where the DataFC value is 0. Two red circles highlight the DataFC values in the fourth and fifth frames (1406 and 1407), which are both set to 4112, indicating a credit error.

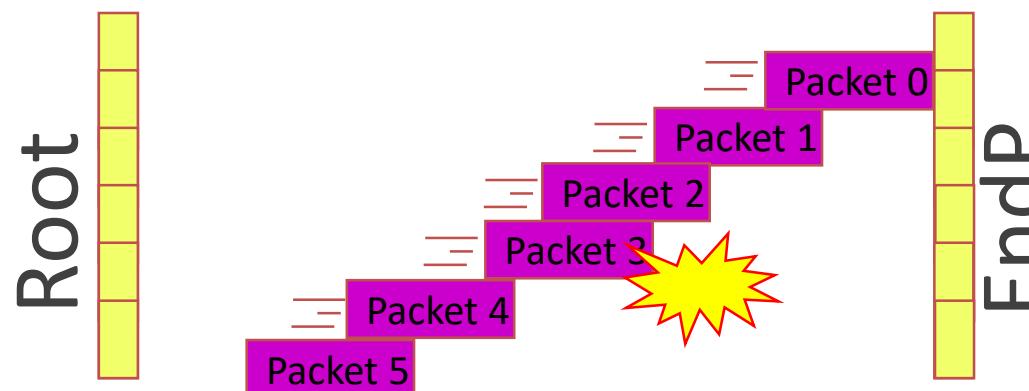
# Data Link Layer Problems

- **Flow control overflow**
- **Why is a packet getting a NAK?**
  - Set up trigger on NAK and look what happened before
- **Credit Limited?**

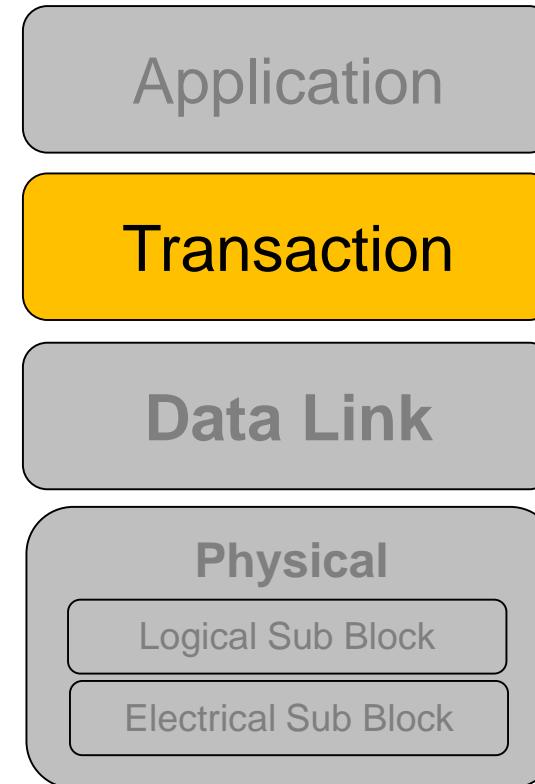
# Test Your Flow Control Knowledge...

## Scenario:

- Root port has 6 credits available for the channel allowing it to send 6 packets to an endpoint.
- The Endpoint acknowledges the first 3 but indicates the fourth packet (seq=3) has failed
  - sends a NAK Seq=2.
- How many credits have been used/consumed?
  - 6? Because 6 packets were sent?
  - 3? Because 3 packets were good and the fourth was Nak'd?
  - 4? 3 good + 1 bad but everything after the bad one is ignored?
  - 0? Because it was NAK'd and all 6 must be resent?



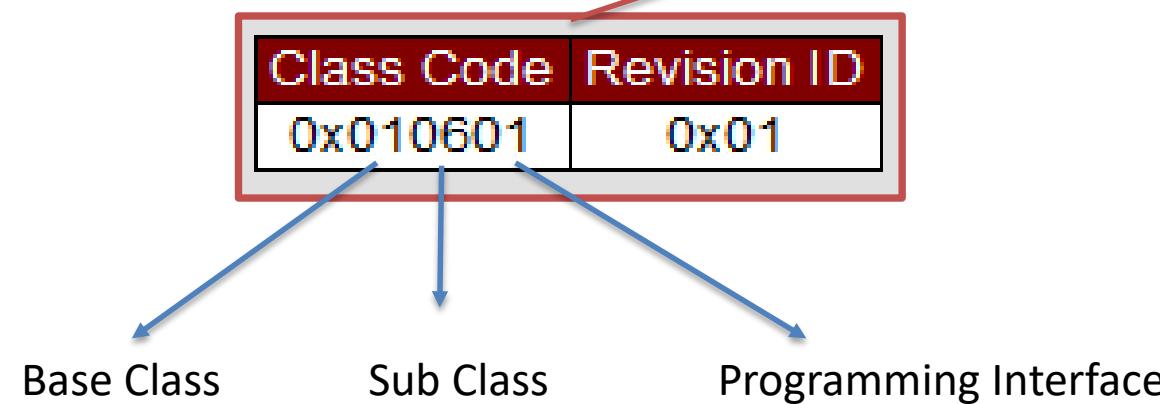
# Class Code Issues



# Class Code

- Class Code used to indicate type of device

Split Tra 110	R→ 5.0 x1	Cfg	CfgRd0	RequesterID 000:00:0	CompleterID 002:00:0	Tag 0	TC 0	VC ID 0	DeviceID 002:00:0	Register 0x008	Status SC	Class Code 0x010601	Revision ID 0x01	Metrics	# LinkTrans 2	Time Delta 1.256 us	Time Stamp 0006. 122 324 510 s
------------------	--------------	-----	--------	-------------------------	-------------------------	----------	---------	------------	----------------------	-------------------	--------------	------------------------	---------------------	---------	------------------	------------------------	-----------------------------------



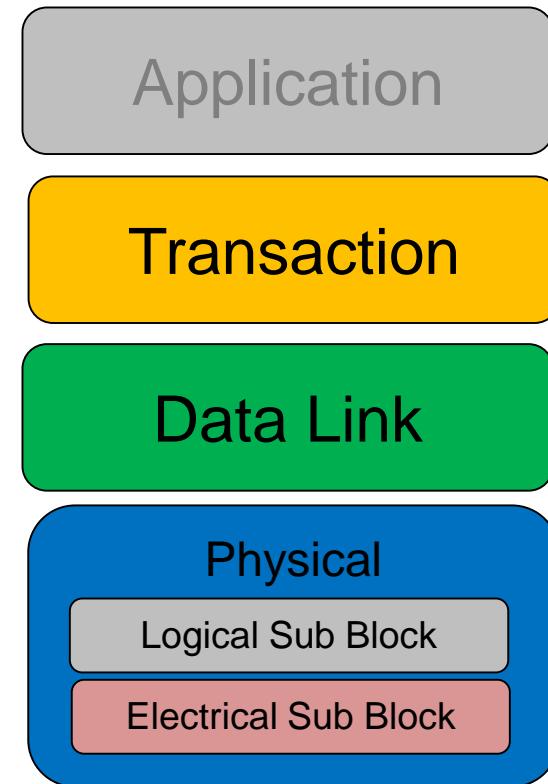
# Class Code Issues



- **Incorrect Class Code assignments cause problems**
- **Showing the device as the wrong type of device can cause the system to be unstable or crash**
- **Examples**
  - BIOS does not support it
  - Incorrect driver is used

Base Class	Sub-Class	Programming Interface	Meaning
01h	05h	20h	ATA controller with ADMA interface - single stepping (see Note 2)
		30h	ATA controller with ADMA interface - continuous operation (see Note 2)
	06h	00h	Serial ATA controller - vendor-specific interface
		01h	Serial ATA controller - AHCI interface (see note 7)
		02h	Serial Storage Bus Interface
	07h	00h	Serial Attached SCSI (SAS) controller - vendor-specific interface
		01h	Obsolete
	08h	00h	Non-volatile memory subsystem - vendor-specific interface
		01h	Non-volatile memory subsystem - NVMHCI interface (see note 8)
		02h	Non-volatile memory subsystem - NVMe interface (see Note 6)
	09h	00h	Universal Flash Storage (UFS) controller - vendor-specific interface
		01h	Universal Flash Storage (UFS) controller - Universal Flash Storage Host Controller Interface (UFSHCl) (see Note 5)
	80h	00h	Other mass storage controller - vendor-specific interface

# Compliance Test Failures



# “PCIe 1.0, 2.0, 3.0, 4.0????”

- **What is the difference?**
  - These are different revisions of the specification
  - They do not necessarily mean a device can support higher speeds
  - The changes go beyond the speed capability and include ECNs and other updates
- **The is a Tendency to incorrectly use Spec version “1.0, 2.0, 3.0, 4.0” synonymously with link rate 2.5GT/s, 5GT/s, 8GT/s and 16GT/s**
  - A PCIe 1.0a or 1.1 device supports 2.5GT/s
  - A PCIe 2.0 device must support 2.5GT/s and can support 5GT/s
  - A PCIe 3.0 device must support 2.5GT/s and can support 5GT/s and/or 8GT/s
  - A PCIe 4.0 device must support 2.5GT/s and can support 5GT/s, 8GT/s and 16GT/s
  - A PCIe 5.0 device must support 2.5GT/s and can support 5GT/s, 8GT/s, 16GT/s, 32GT/s

# “PCIe 1.0, 2.0, 3.0, 3.1, 4.0, 5.0???”

- **The most significant change in specification revisions is to add higher speed capabilities**
- **However, there are many other ECNs that were approved after the release of a specification that are rolled into newer releases**
- **For example, there is a 3.1 revision of the specification that included ECNs proposed after the release of the 3.0 specification**
- **The PCI Express 5.0 Base Specification is now the latest specification for ALL speeds.**

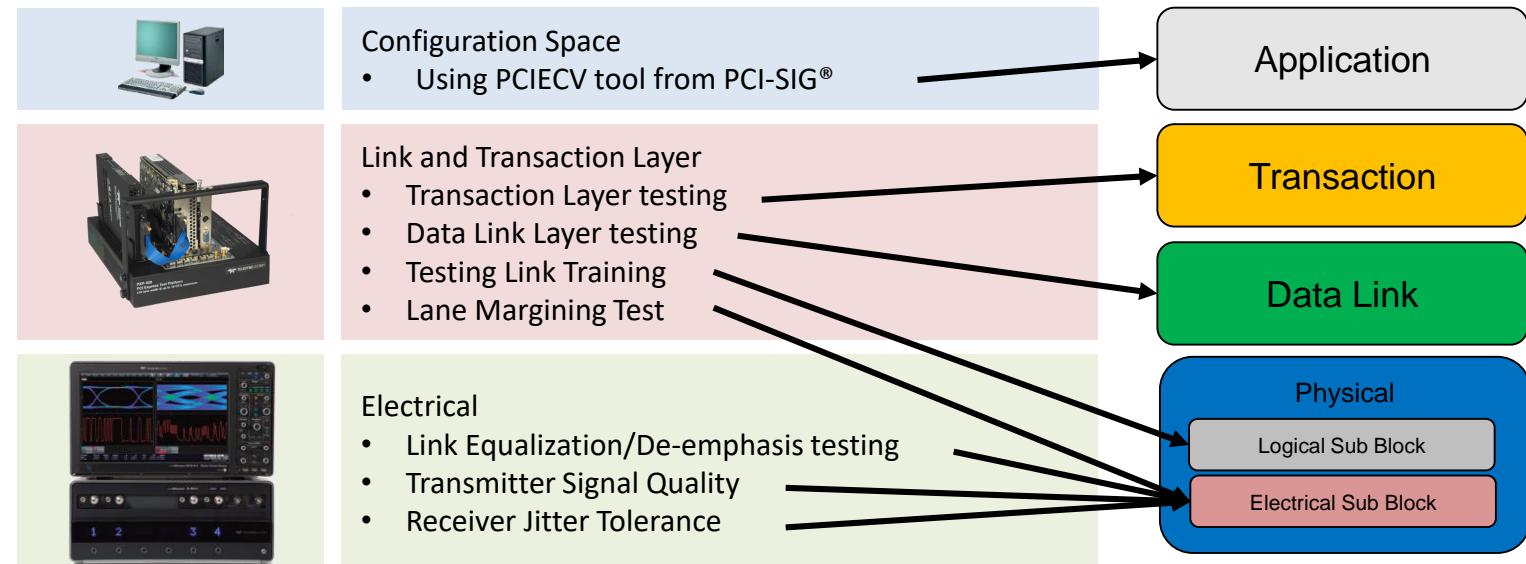
# “PCIe 1.0, 2.0, 3.0, 4.0????”

- **There are many changes in Configuration Space that can results in differing test results**
- **For Example: The way PCI-SIG tests devices is not the same for a PCIe 1.1 device and a PCIe 3.0 device that only supports 2.5GT/s**

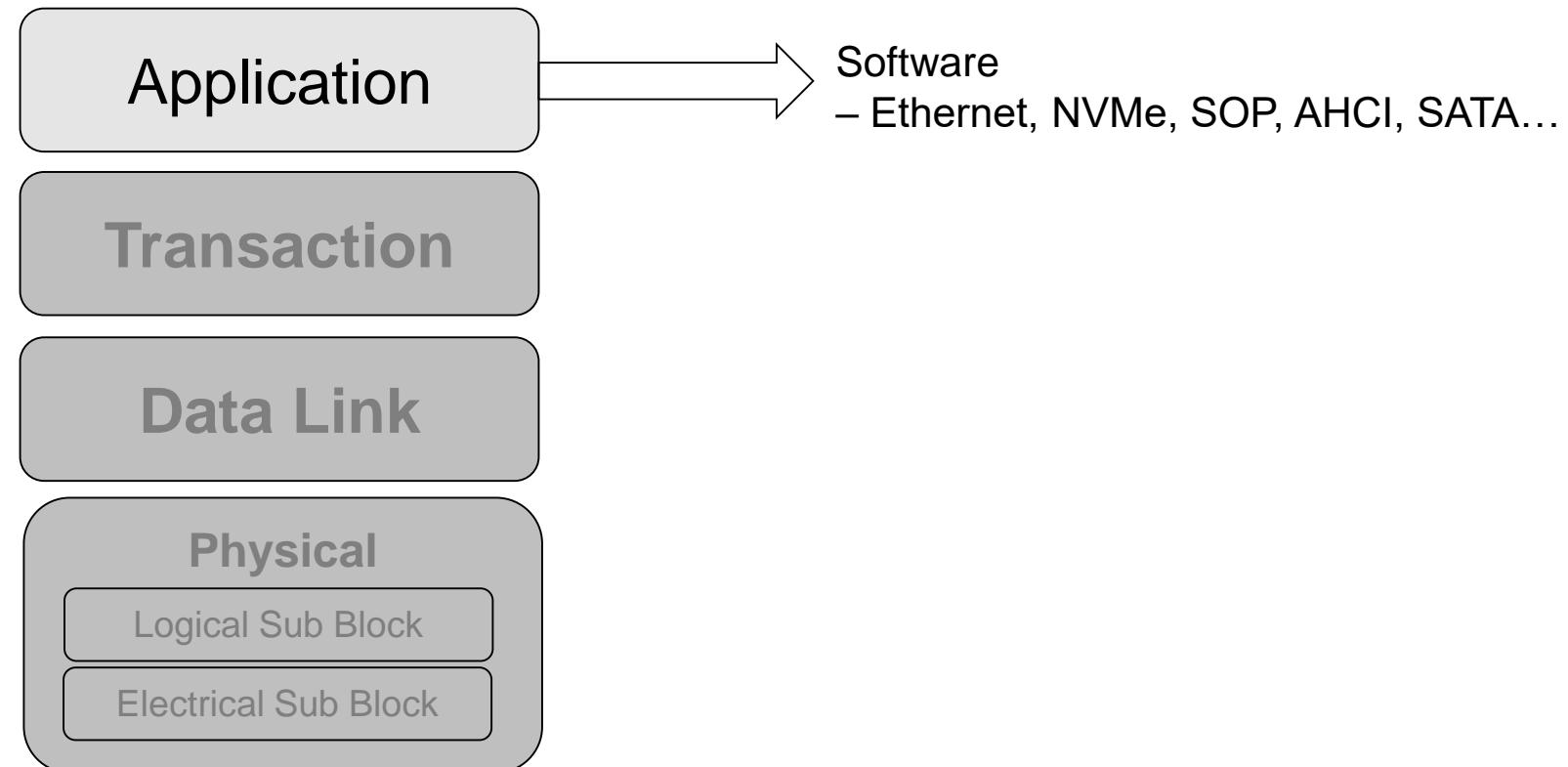
# PCI-SIG® Compliance Testing



- How Does this fit into the layered model?
- Compliance Testing is split out into sections:



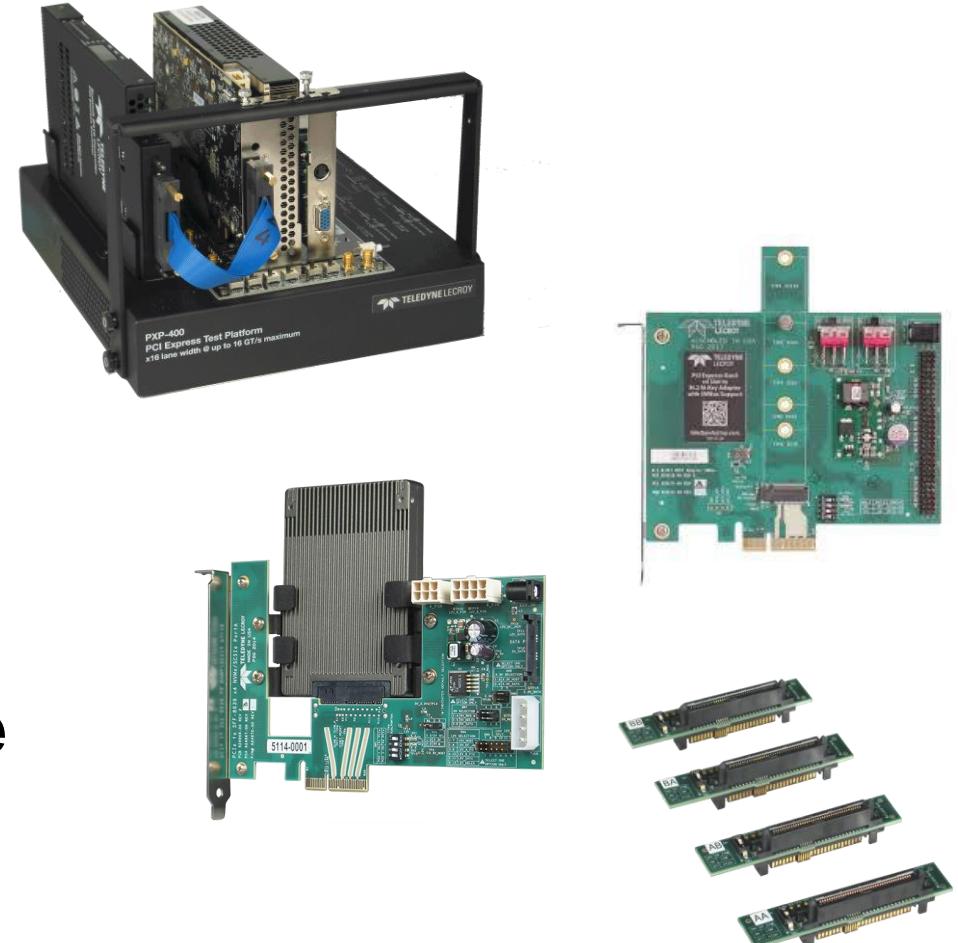
# PCI Express Storage



# PCIe Storage Devices – U.2, U.3, M.2



- **Where does this fit into the layered model?**
- **How do I debug?**
- **“Do I care about PCI Express!!!”**
  - Still have PCIe interfaces
  - Adapters allow compliance test
- **Compliance Programs for Storage protocols like NVMe are outside of the scope of PCI-SIG, however PCI-SIG is impacted by form factor changes**



# How Do I Debug These Products?

- **It may take a different way of thinking to use a PCI Express analyzer for debugging a storage problem**
- **The storage protocols reside within the PCI Express payload data, and requires further decoding**
- **It is possible that the only interface to these devices is PCI Express**
- **Other form factors are becoming very common; U.2, M.2 for example**

# NVMe Example

Teledyne LeCroy PCIe Protocol Analysis - [C:\Users\Public\Documents\LeCroy\PCIe Protocol Suite\Sample Files\NVMe_Z3DriveEmulation.pex]																		
Trace View																		
NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM			
21		Read	0x0001	0x0001	0x0003	128 dwords	00000000:00000000	00000002:2DD9B000	00000000:00000000	00000000:00000000	0x0000	0x0	0	0	0			
NVMe	H	Device ID	QID	SQyTDBL	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp								
120		006:00:0	0x0001	0x0004	0x0004	NVMeLeCroy000000	1											
NVMe	H	Device ID	QID	CID	Address	IOSQ	OPC	FUSE	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM		
121		006:00:0	0x0001	0x0003	00000002:2E0830C0	Read	Normal op	0x0003	0x0003	0x00000001	0x0000	0x0	0	0	0	0		
Split Tra	R+	2.5	Mem	MRd(64)	RequesterD	CompleterD	Tag	TC	VC ID	Address	Status	Data	Metrics	# LinkTrans	Time Delta	Time Stamp		
878	x1			001:00000	006:00:0	000:00:0	21	0	0	00000002:2E0830C0	SC	16 dwords	2	372.000 ns	0039			
Link Tra	R+	2.5	TLP	MRd(64)	Length	RequesterD	Tag	Address	1st BE	Last BE	VC ID	Explicit ACK	Metrics	# Packets	Time Delta	Time Stamp		
1928	x1			998	Mem	001:00000	16	006:00:0	21	00000002:2E0830C0	1111	1111	0	Packet #377119	2	372.000 ns	0039	
Link Tra	R+	2.5	TLP	Cpl	Length	RequesterD	Tag	CompleterD	Status	BCM	Byte Cnt	Lwr Addr	Data	VC ID	Explicit ACK			
1929	x1			930	Cpl	010:01010	16	006:00:0	21	000:00:0	SC	0	64	0x40	16 dwords	0	Packet #377121	
NVMe	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp					
122		006:00:0	0x0001	0x0003	00000002:2DD9B000	0x00000080	128 dwords	NVMeLeCroy000000	1	4		107.984 us	0039 . 509 627 694					
NVMe	D	Device ID	QID	CID	Address	IOCO	SQHD	SQID	CID	P	DW0	Reserved	ST	SCT	SC	M	DNR	
123		006:00:0	0x0001	0x0003	00000002:2E093030	0x0004	0x0001	0x0003	1	0x00000000	0x00000000	0x00000000	Generic Command Status	Successful Completion	0	0	NVMe	
NVMe	D	Device ID	QID	CID	Address	Interrupt	Type	Vector	Message	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp				
124		006:00:0	0x0001	0x0003	00000000:FEE3F00C	MSI	1	0x000049A9	NVMeLeCroy000000	1	1		29.116 us	0039 . 509 757 326 0 s				
NVMe	H	Device ID	QID	CQyHDBL	IO COH	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp								
125		006:00:0	0x0001	0x0004	NVMeLeCroy000000	1	10.272 us	0039 . 509 786 442 0 s										
NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM			
22		Read	0x0001	0x0001	0x0004	128 dwords	00000000:00000000	00000002:2DFC3CC0	00000000:00000000	00000000:00000000	0x0000	0x0	0	0	0			
NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM			
23		Read	0x0001	0x0001	0x0005	128 dwords	00000000:00000000	00000002:2DFCD980	00000000:00000000	00000000:00000000	0x0000	0x0	0	0	0			
NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM			
24		Read	0x0001	0x0001	0x0006	128 dwords	00000000:00000000	00000002:2DD9D000	00000000:00000000	00000000:00000000	0x0000	0x0	0	0	0			
NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM			
S → E QuickTiming markers not set																		
Summit T48 SN:2000 US — ● ● ● ● ● DS — ● ● ● ● ●																		
Ready Errors detected! Search: Fwd																		

The NVMe Commands are actually PCI Express Memory Reads and Writes to specific addresses

- **The underlying PCIe traffic has an impact on performance of the Storage layers**
- **Possibility of multiple devices over one link**
  - Devices behind a switch
  - Virtualized devices
  - Built in scalability of technologies such as NVMe

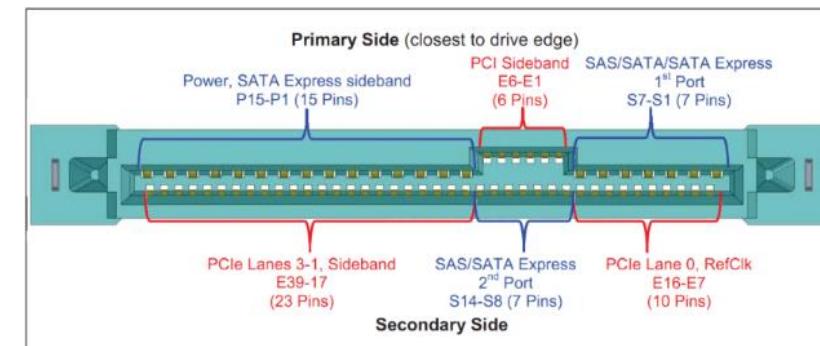
# Device ID Changes

- Something to watch out for when using the some drivers is that the Bus:Device:Function of the device may change during initialization
- It is important when testing NVMe devices to ensure that you are observing the correct Device ID
- This also applies to a link upstream of a switch, there may be multiple streams of traffic on one link.

NVMe	H	Device ID	QID	CID	Address	IOSQ	OPC	FUSE	PSDT	CID	NS
121		006:00:0	0x0001	0x0003	00000002:2E0830C0		Read	Normal operation	PRP	0x0003	0x000
	Split Tra 878	R↔ 2.5 x1	Mem	MRd(64)	RequesterID	CompleterID	Tag	TC	VC ID	Address	
				001:00000	006:00:0	000:00:0	21	0	0	00000002:2E0830	

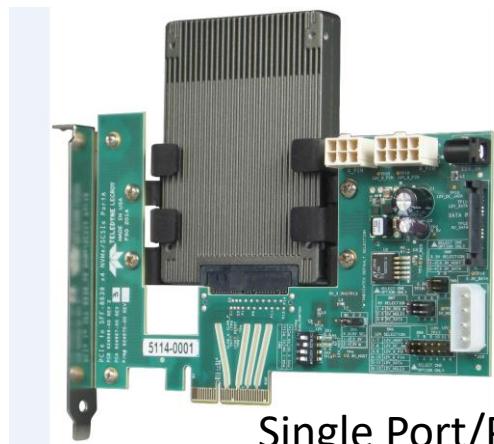
# PCI Express U.2™ Connector

- Used as a “Drive” form factor
- Connector supports PCIe, SAS, SATA
- Supports single or dual link configuration
  - Dual Port Enable pin grounded for slot that supports dual link
  - Single x4 link or dual x2 links

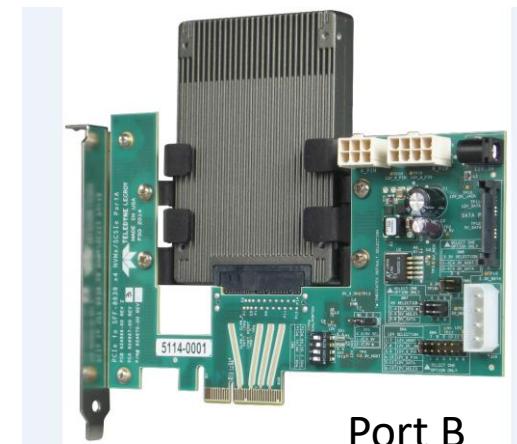


# Dual Port PCI Express U.2

- **Dual Port PCIe/NVMe devices are starting to be developed**
  - A Dual Port device has 2 separate independent PCIe links
  - It requires 2 analyzers and a separate Interposer to debug
- **Compliance Testing requires each port to be tested separately**



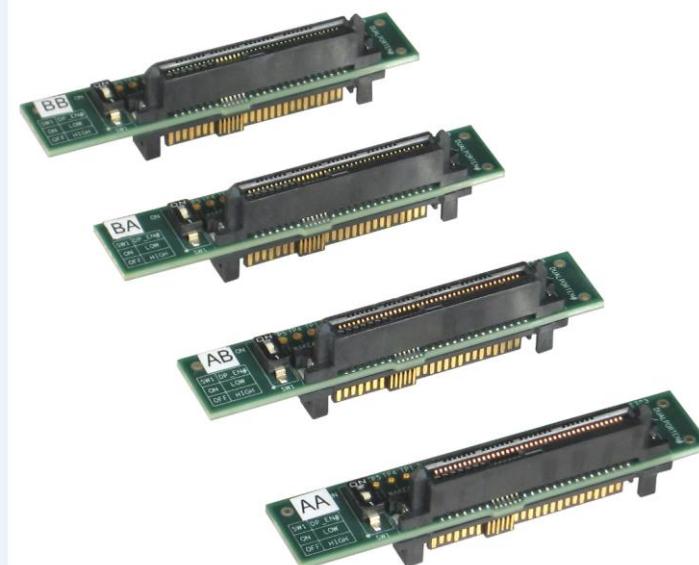
Single Port/Port A



Port B

# Debugging Dual Port

- Try to set up analysis tools separately for each port before trying to look simultaneously
- If necessary, use a transposer to block one or other of the ports to work on a single port at a time



# Summary and Conclusions

- **The easiest way to debug PCI Express problems is to determine which layer is showing the problem**
- **Using the correct tool will make the job much easier**
- **Use a “Snapshot” trigger first of all to assess what is going on**
- **Ensure lower level (electrical) layers are compliant and within spec before debugging protocol problems**
- **Remember the PCI-SIG compliance testing is not a substitute for thorough validation, it is a useful sanity check**

**Thank you for attending the  
PCI-SIG Developers Conference  
Asia-Pacific Tour 2019.**

**For more information please go to**  
**[www.pcisig.com](http://www.pcisig.com)**

