



# Reliable Data Transmission Features of PCI Express®

Gord Caruk, Fellow  
AMD

Presented by  
Betty Luk, Member of Technical Staff



# Agenda

- Introduction
- Dissection of a recovered error
- Loss of a single bit
- Loss of a stream of bits from a lane
- Catastrophic loss of one or more lanes
- Demo

# Introduction

- In the 80s and early 90s, ISA, PCI, and others, were the predominant interfaces in the PC
- These interfaces were relatively simple
  - ✓ Protocol complexity
  - ✓ Total number of transistors to implement
  - ✓ No complex analog circuitry
  - ✓ I/O's were simple
  - ✓ Power distribution, crosstalk were typical issues
- Serial interfaces, for eg. PCIe<sup>®</sup>, are much more complex
- Statistical failures are more prevalent as frequency increases

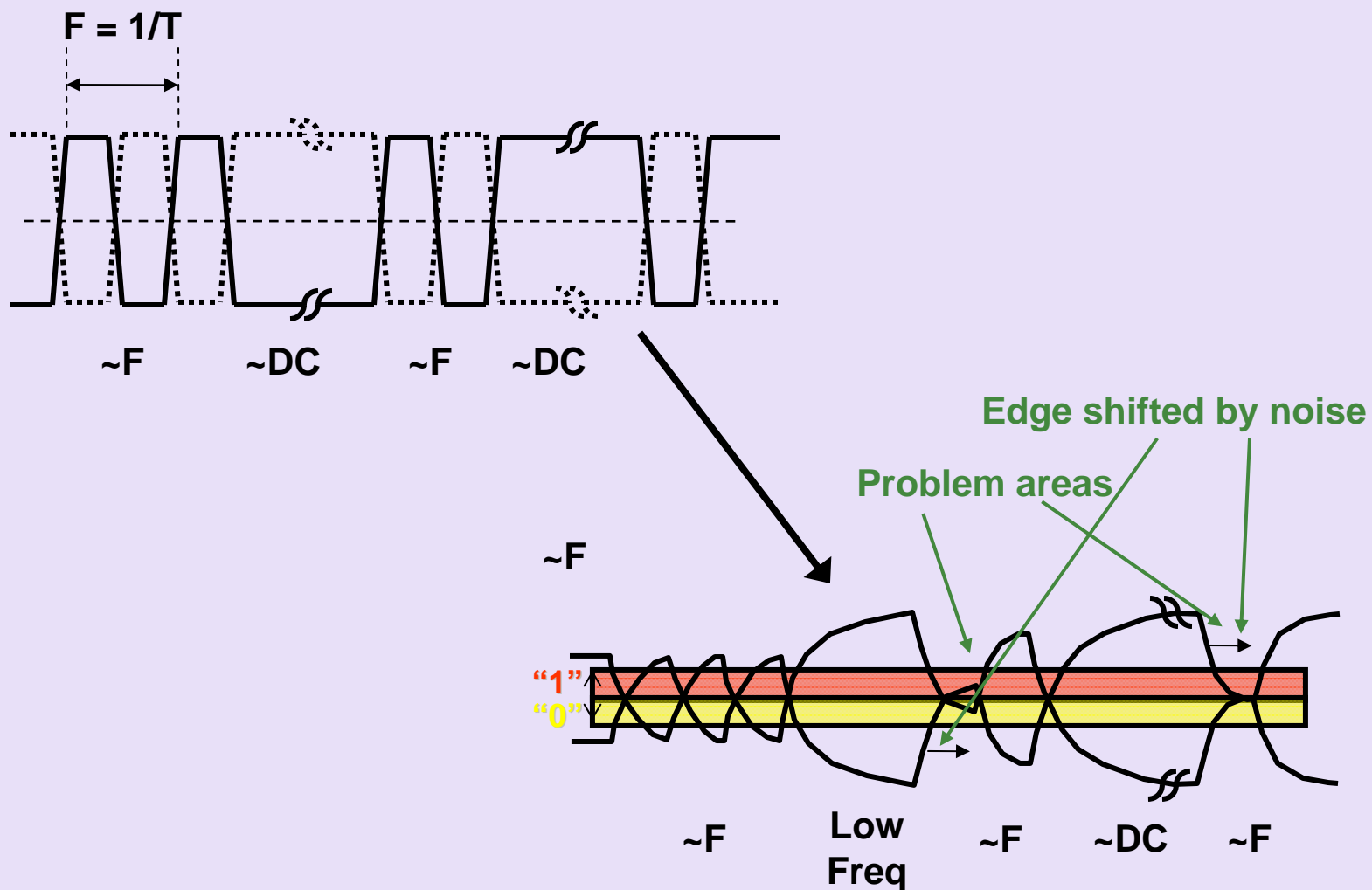
# Introduction

- In the 80s and early 90s, the most common causes of blue screen were from software
- Hardware failures on the connection between components were very low in comparison to the sum total of all failures
- Software and other components are doing better, and are closing the gap
- Mechanisms exist in PCIe to enhance its ability to tolerate faults and reliably transfer information from one component to another

# Recovered Error Dissection

- Required for a error recovery to be occur:
  - ✓ An error or failure
  - ✓ Detecting the error
  - ✓ A mechanism to recover from the error
  - ✓ Invoking that mechanism

# Single Bit Error



# 8b/10b Decode Error

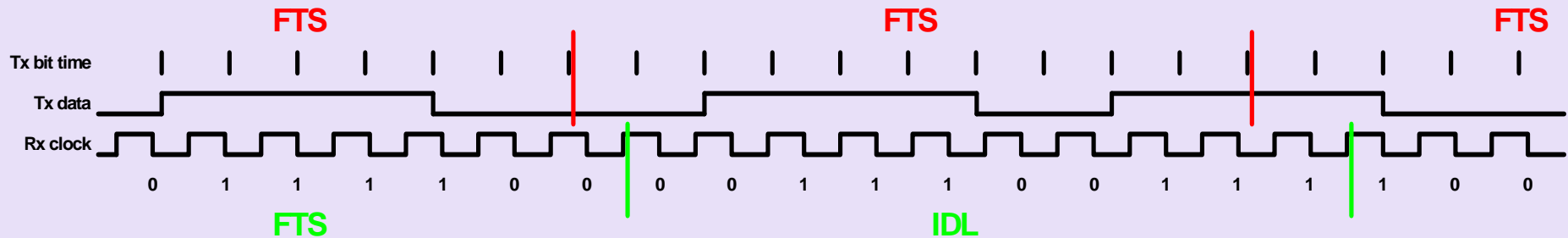
- 10 bit data has 1024 combinations
- 464 of these are legitimate
- 560 cause a decode error
- Single bit error MAY cause an 8b/10b decode error
- SBE may cause a disparity error
- SBE may go undetected by the 8b/10b decoder

# CRC

- CRC covers SBE in:
  - ✓ Data payload
  - ✓ Header
  - ✓ CRC
  - ✓ Sequence number
  - ✓ END symbol, but with less accuracy
- CRC does not detect SBE in:
  - ✓ STP symbol
- Detected CRC error generates a Nak
- Nak forces the TLP to be replayed



# Loss of Symbol Lock



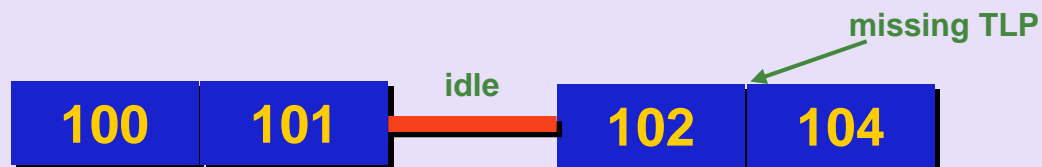
- Caused by excessive jitter, improper B/W in receiver CDR
- Causes ALL successive symbols to be bad until the next symbol locking opportunity
- Lose one lane of data for a long time
- Replay will happen quickly, but receiver may not have recovered

# Ack Latency Timeout

- All TLPs receive an Acknowledgement
- Ack is expected by the Requester
- Transmitter check ensures the receiver is seeing TLPs
- If an Ack is not received, Ack latency timeout occurs, causing replay
- Some implementations do not reacquire symbol lock in L0
  - ✓ Stuck replaying
- Replay num rollover will cause transition to Recovery where (hopefully) a robust reacquire of symbol lock will occur
- Is possible to lose Acks without any loss in TLP transmission reliability

# Sequence numbers

- An incrementing number is prepended to every new TLP
- The receiver keeps an expected next sequence number, so it knows what to expect in a received TLP
- Receiver checks the number received against the expected sequence number
- Detects one or many lost TLPs
- Receiver check ensures TLPs are not lost and stay in order
- Results in Nak, replay, replay num rollover



# Structured Packet Definition

	+0								+1								+2								+3							
	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
Byte 0 >	R	Fmt x 0		Type				R	TC		Reserved				T D	E P	Attr		AT		Length											
Byte 4 >	Requester ID								Tag								Last DW BE				1st DW BE											
Byte 8 >	Address[31:2]																								R							

- All packets follow a standard definition
- Checked that they are framed correctly
- Fields are interrogated to determine they are populated properly

# Handling with DLLPs

- DLLP do not have inherent handshake
- Resolved by resending DLLPs
  - ✓ Lose of FCC update
    - A 'heartbeat' FCCU is sent every 30 usec
  - ✓ Loss of Ack
    - Subsequent Ack/Nak
    - Fundamental component of the sequence number / Ack handshake – will initiate replay

# Down-configuration

- One or more lanes may stop working
  - ✓ Result of unique logic per lane
  - ✓ Or common logic used across several lanes
- Link will transition to Recovery, then Configuration to reduce link width
- Link gets back to L0 while maintaining LinkUp
- LinkDown is BAD

## Ack replay timer overflow in RC

## TS1s on all lanes

## TSx1 on some lanes

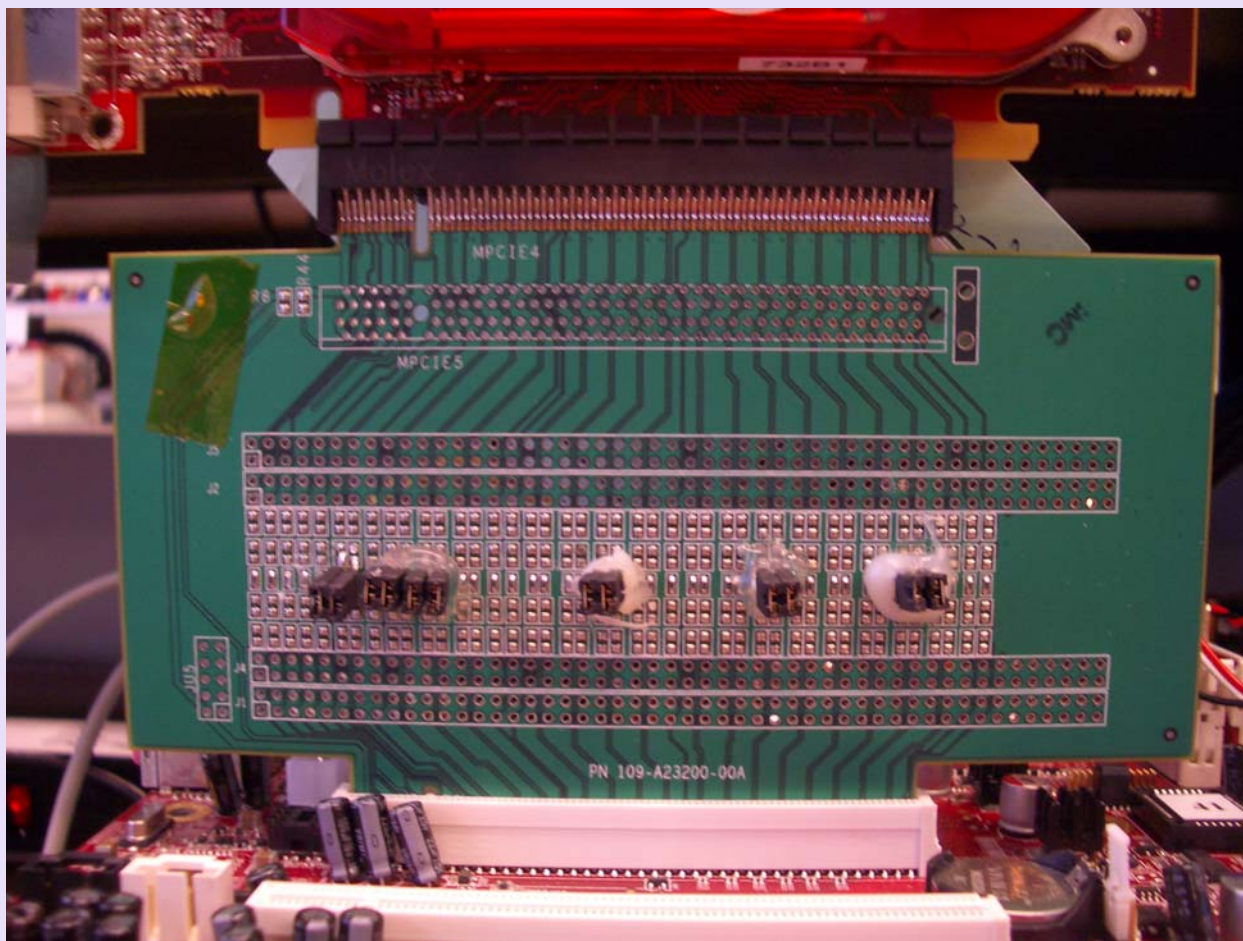


# DEMO

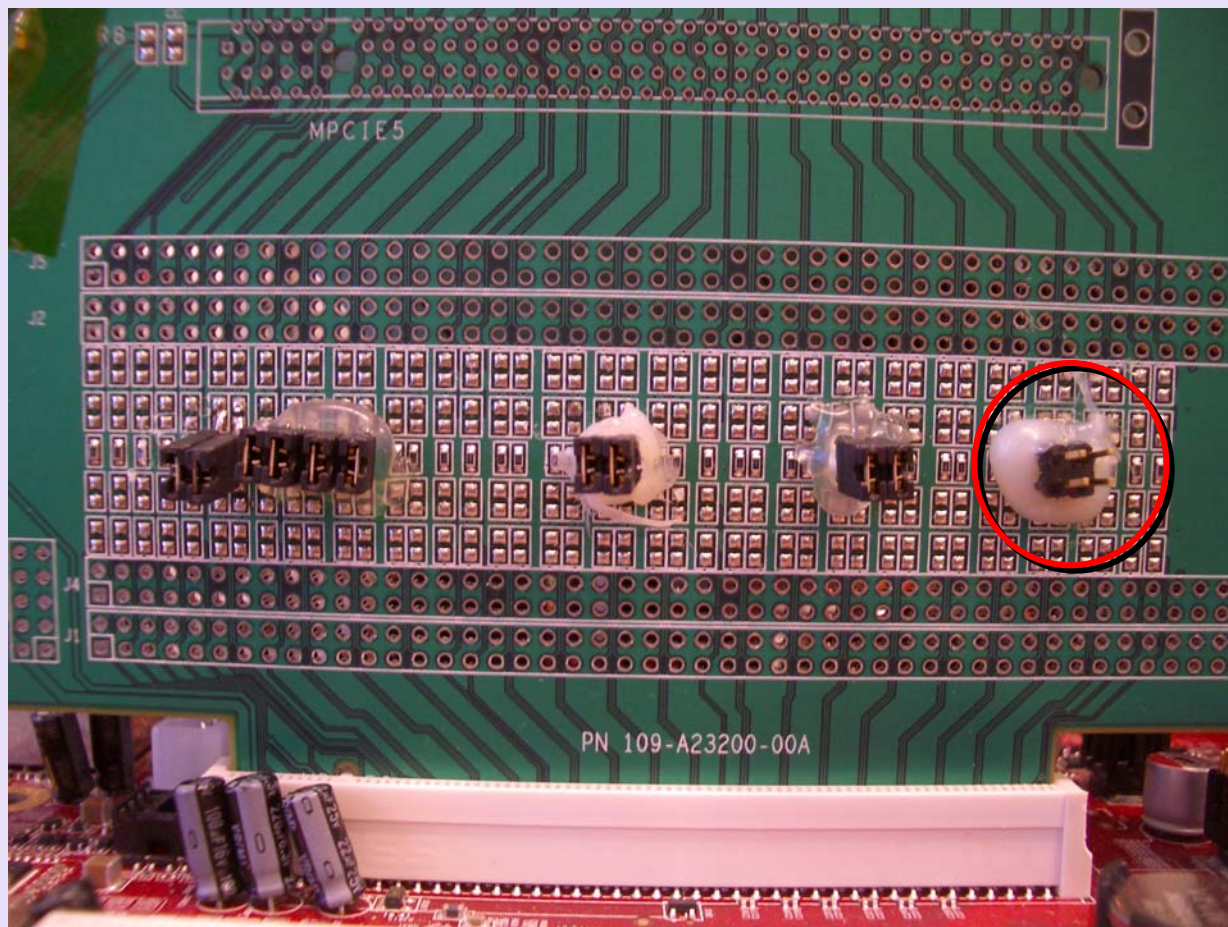
- Demonstration of link down-configure to stay at LinkUp even though lanes keep breaking
- A x16 “pass-through” riser card is used
- Lanes 0, 1, 2, 6, 10, and 14 (upstream lanes) are “jumped”
- Lane is disconnected to downstream Rx when jumper is removed



# Demo Setup



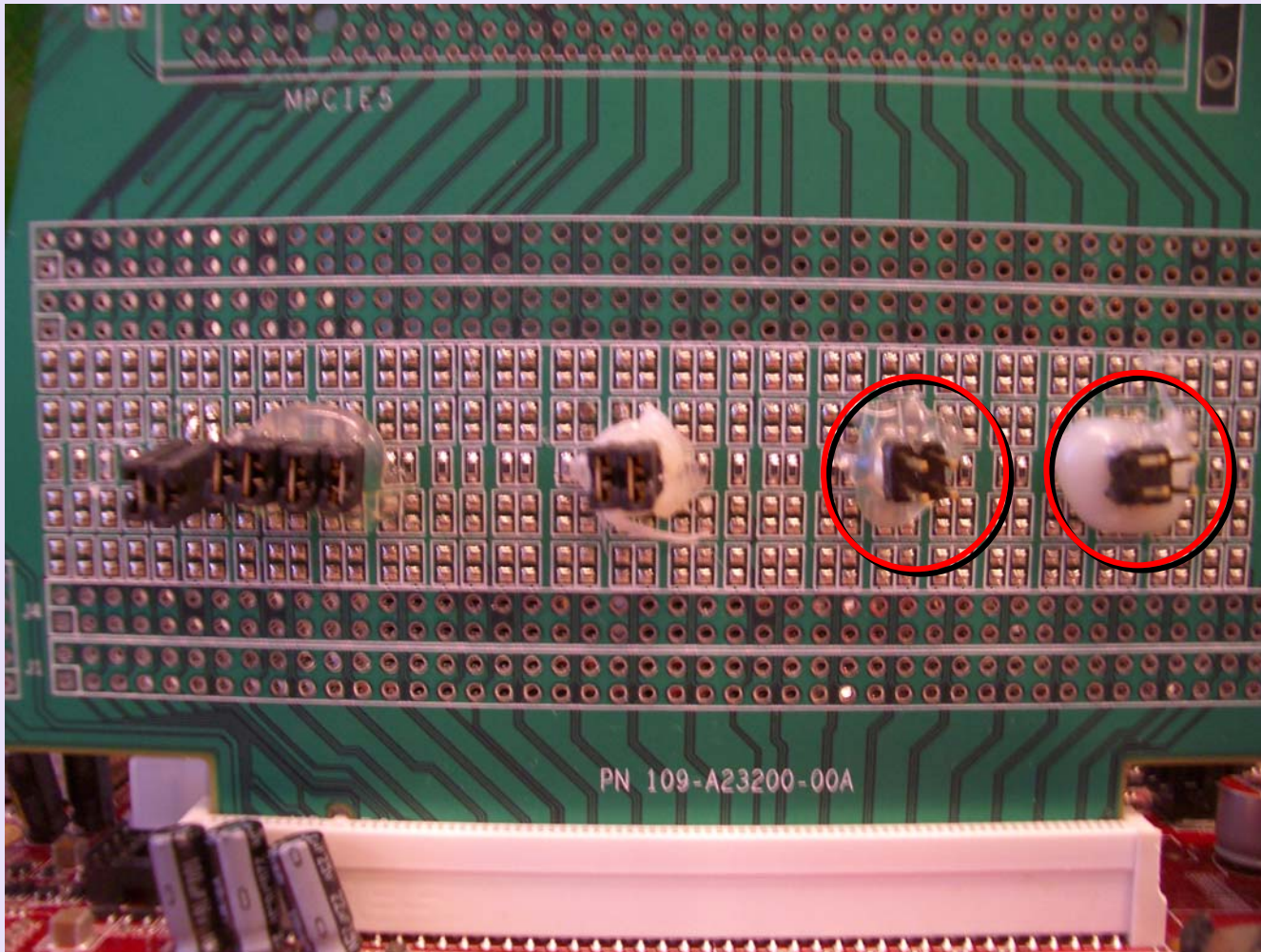
# Disconnect Lane 14



- Link is operating at x12

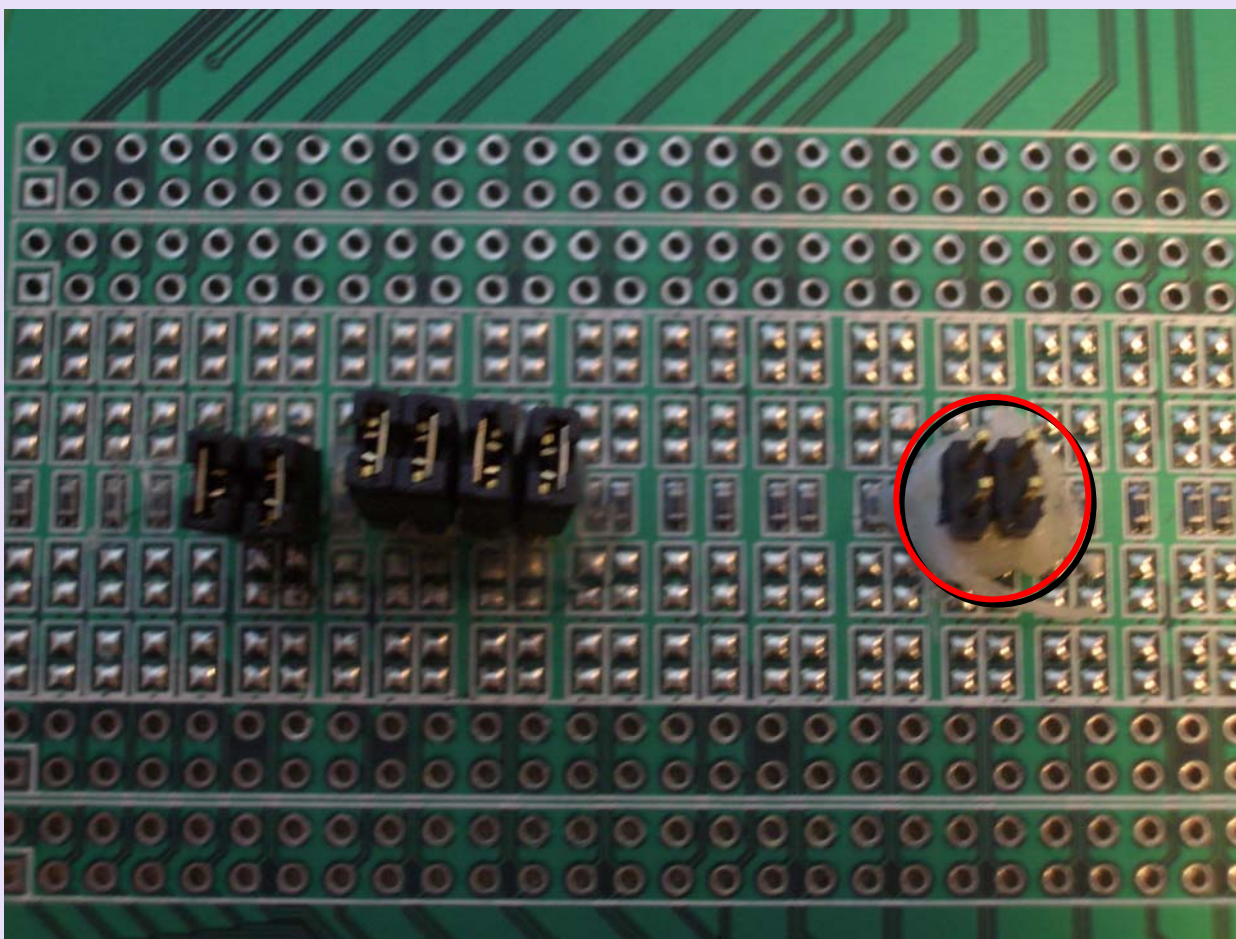


# Disconnect Lane 10



- Link is operating at x8

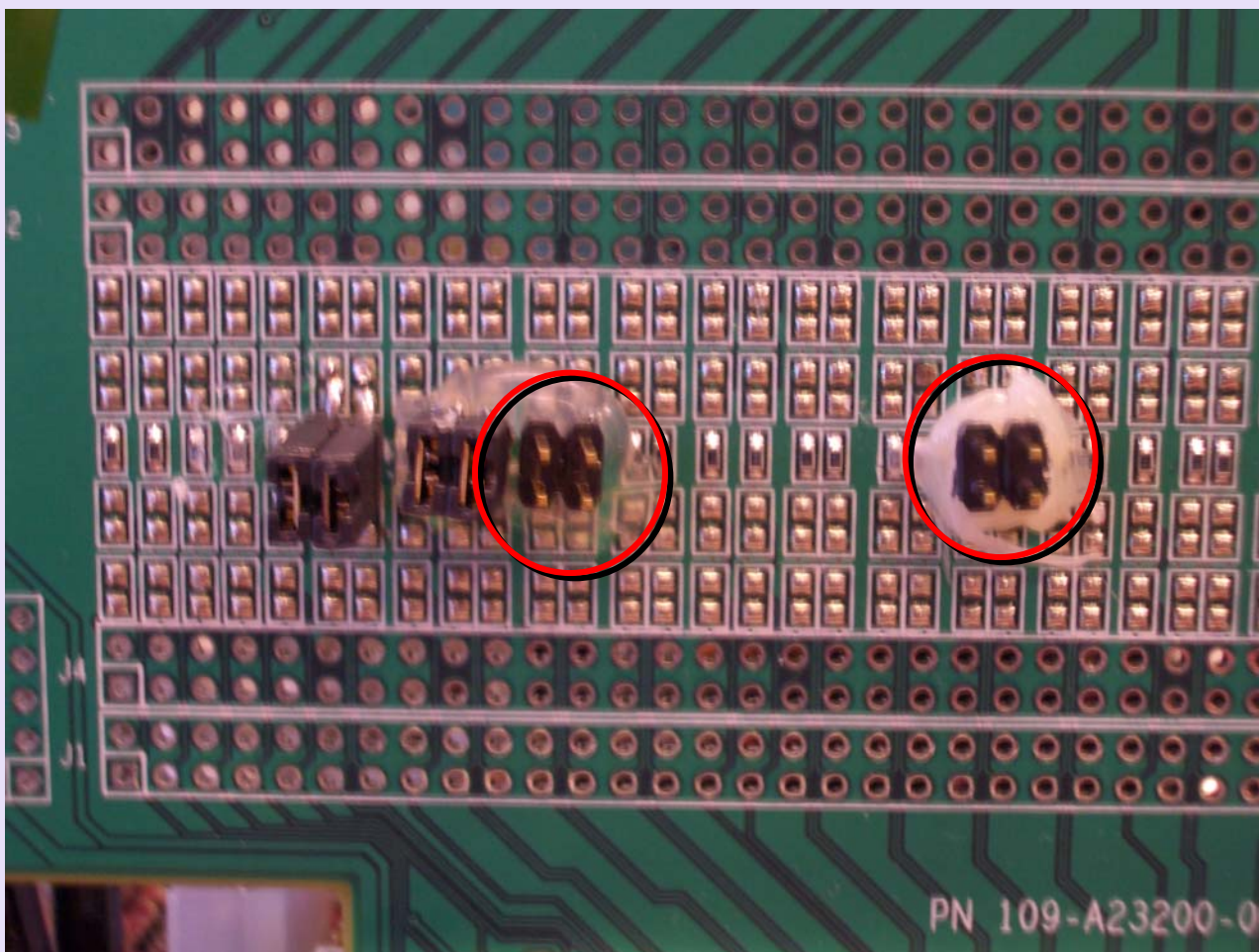
# Disconnect Lane 6



- Link is operating at x4

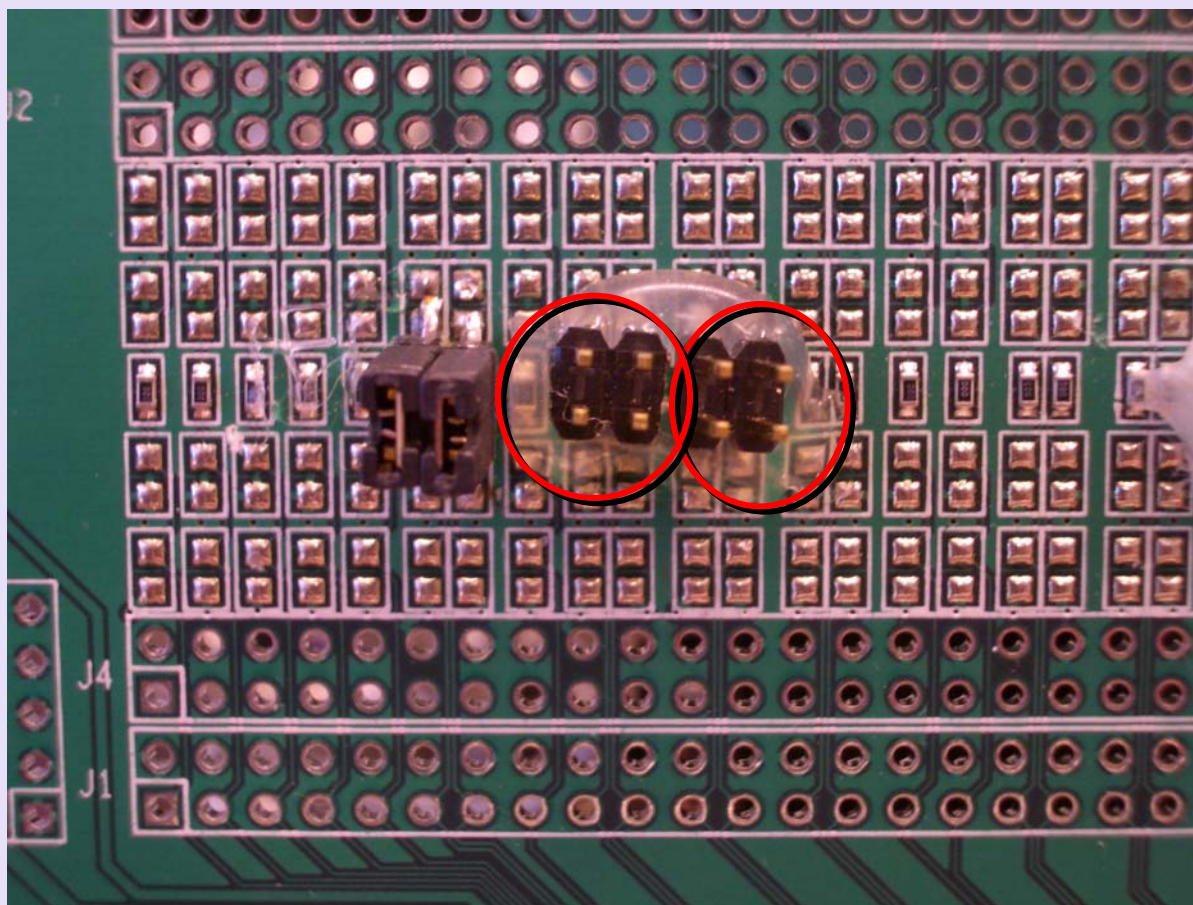


# Disconnect Lane 2



- Link is operating at x2

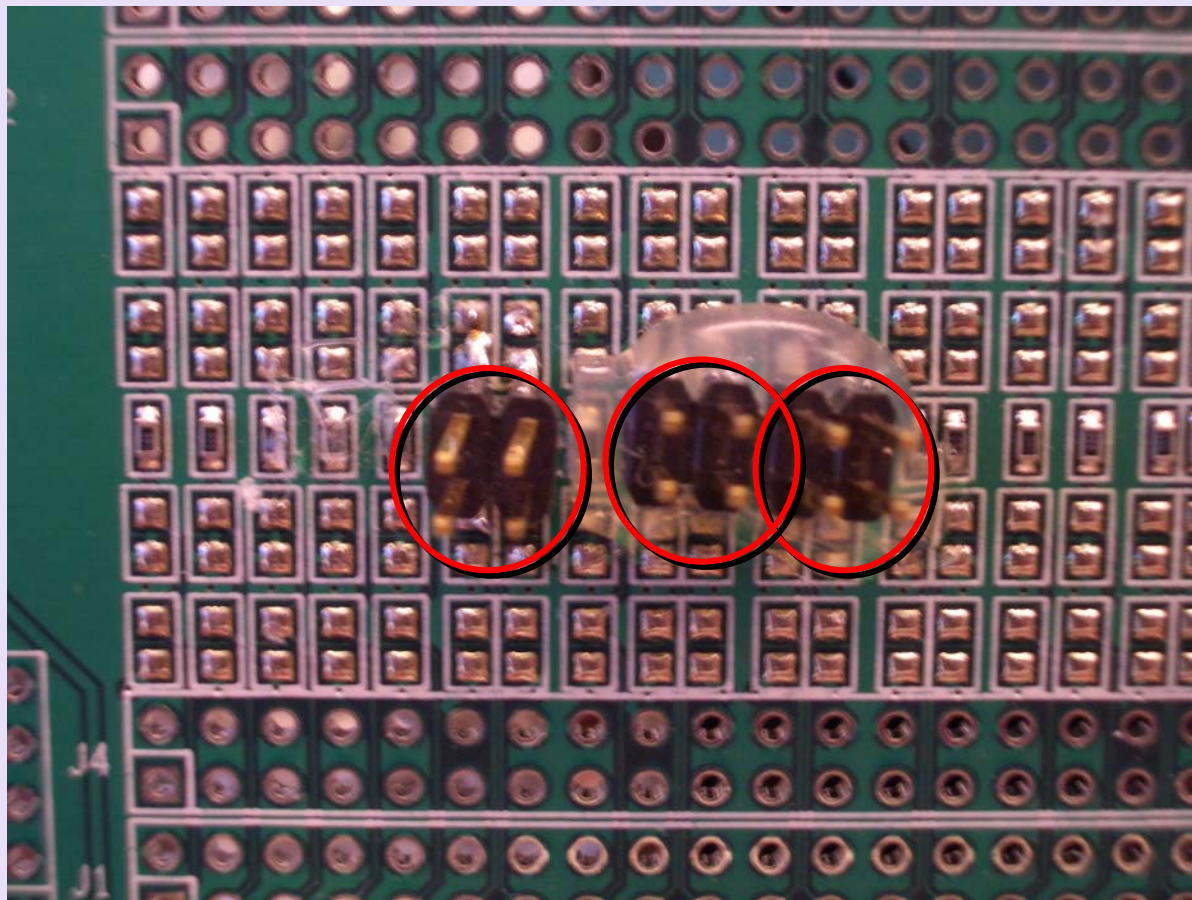
# Disconnect Lane 1



- Link is operating at x1

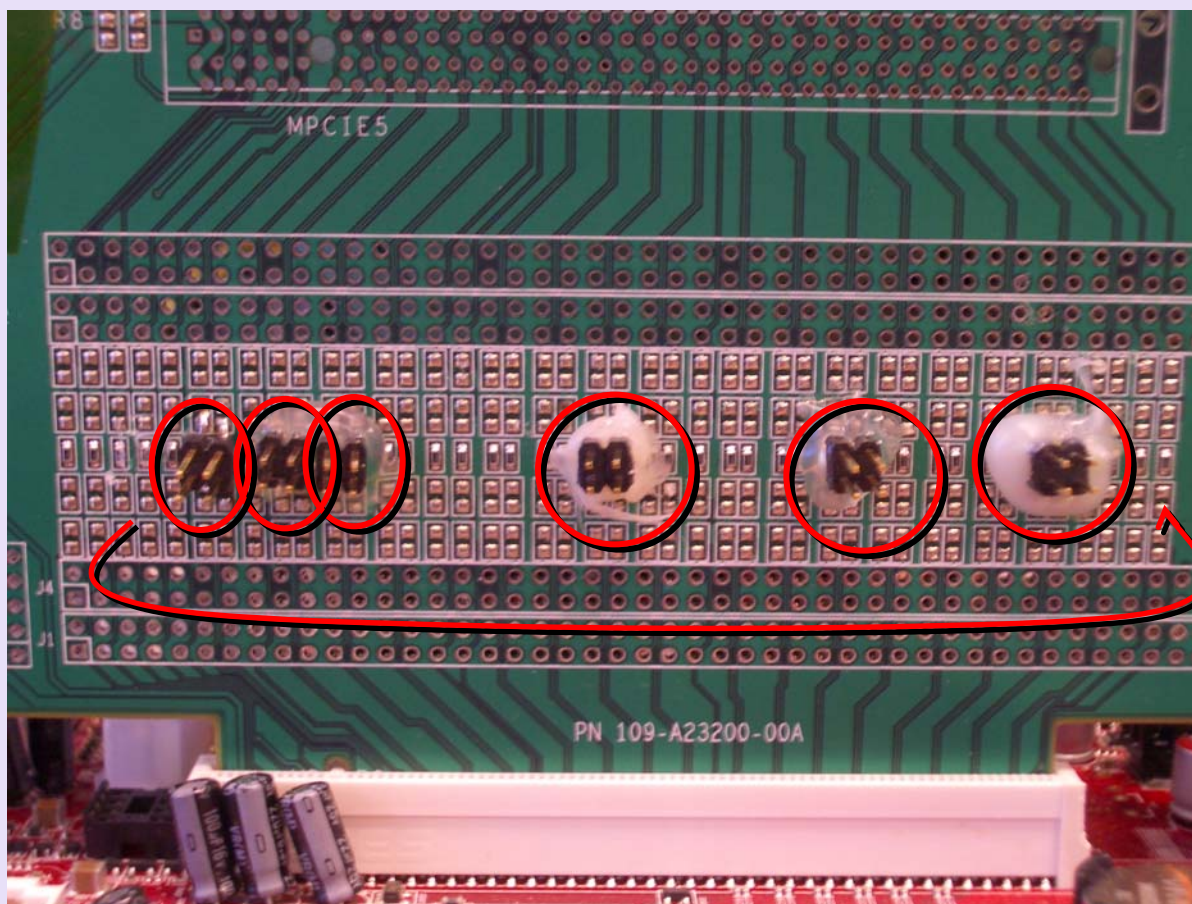


# Disconnect Lane 0



- Link is operating at x1 (lane reversed)

# Link up using lane 15



- Lanes 14, 10, 6, 2, 1, 0 were disconnected in order



# Summary

- Fault Cases
  - ✓ Single bit errors
  - ✓ Loss of symbol lock
  - ✓ Catastrophic loss of a lane
- Detecting errors
  - ✓ 8b/10b decode errors
  - ✓ CRC
  - ✓ Ack latency timeout
  - ✓ Sequence numbers
  - ✓ Structured packet definition
  - ✓ Resending DLLPs
- Fault recovery mechanisms
  - ✓ Replay
  - ✓ Retraining via Recovery
  - ✓ Down configuration

Thank you for attending the  
PCI-SIG Developers Conference  
Europe 2007.

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