PCI Express® 1.1 PHY Design Considerations

Milpitas, CA August 21, 2006

Joe Winkles

Staff Architect, MindShare, Inc.

Agenda

- Logical Physical Layer
- Electrical Physical Layer
- Link Training and Initialization (LTSSM)
- Layout Considerations

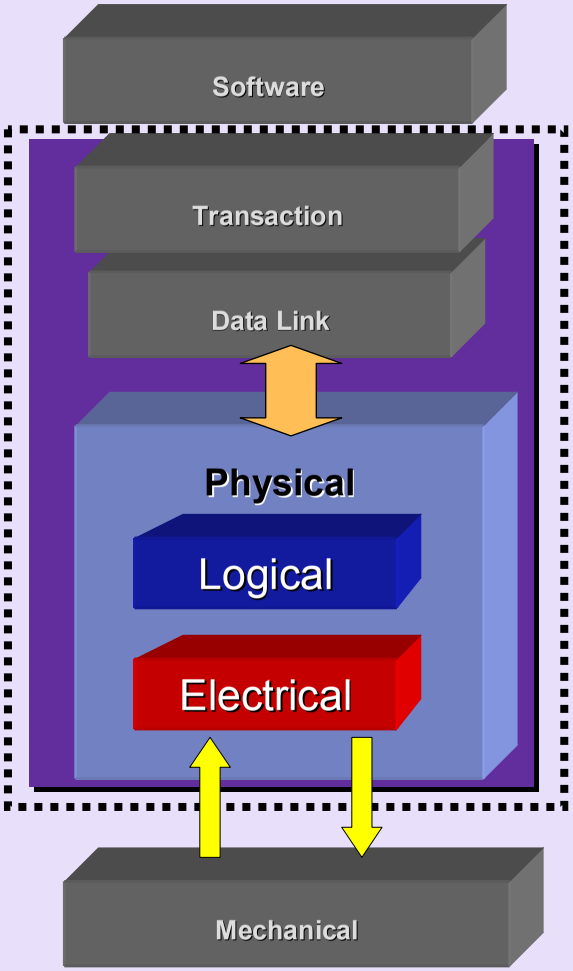
Agenda

- Logical Physical Layer
- Electrical Physical Layer
- Link Training and Initialization (LTSSM)
- Layout Considerations

PCI Express Physical Layer

- **PCI – PHY is digital in nature**
  - ✓ Parallel multi-drop bus
  - ✓ Bits, a clock, setup and hold times…
    - – Jitter essentially ignored
- **PCI Express – PHY is analog in nature**
  - ✓ Based on serial technology
  - ✓ Techniques developed by the communication industry
- **A transition for microprocessor based systems**
  - ✓ Microwave theory/physics challenges dominate
  - ✓ Two PLLs communicating directly

PHY Layer Design Basics

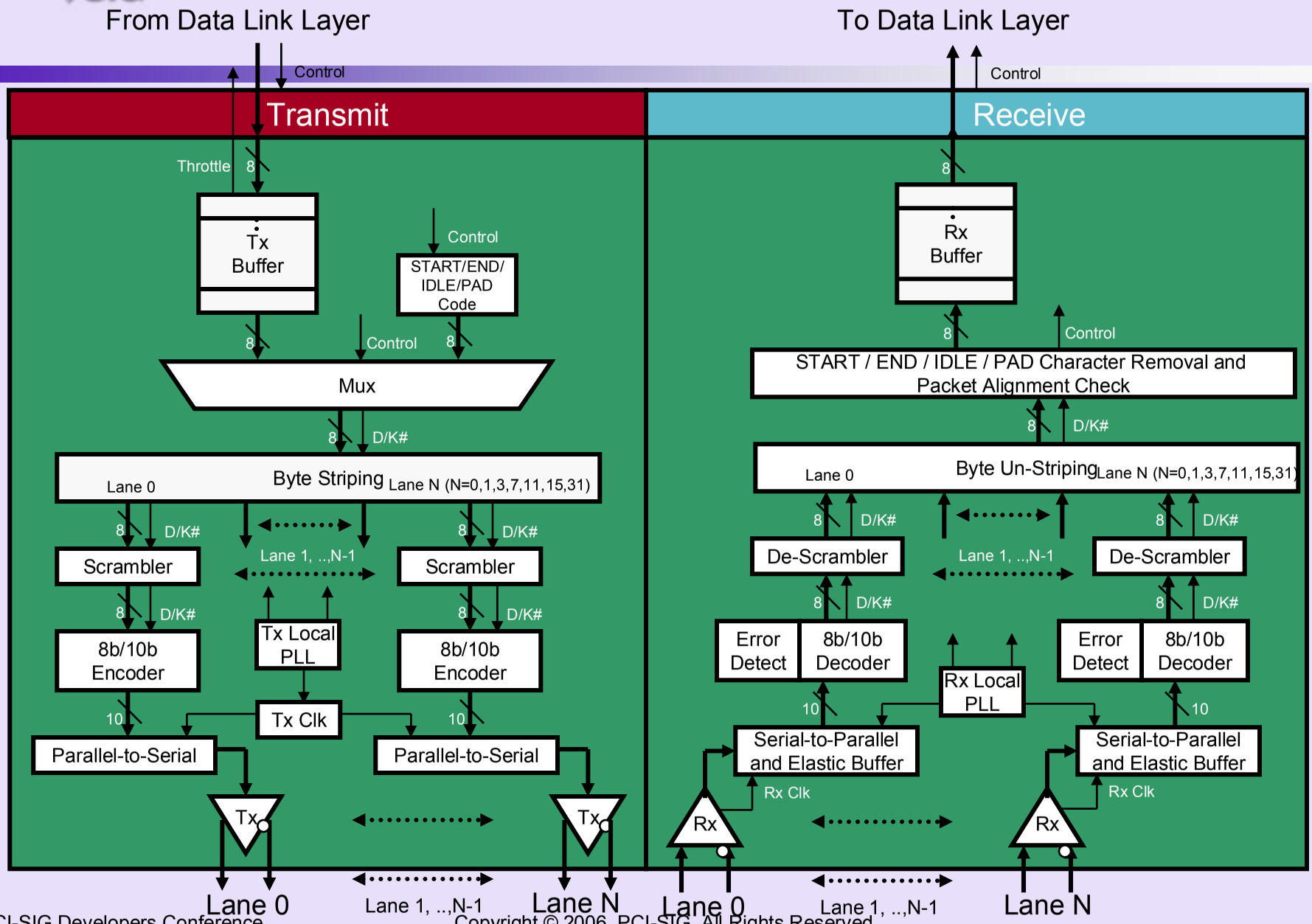


- **Logical Functions**
  - ✓ Encoding/decoding/scrambling
  - ✓ Reset, initialization, de-skew
  - ✓ Built in test modes
  - ✓ Configuration:
    - Speed, link width, lane mapping, polarity
  - ✓ Link power management

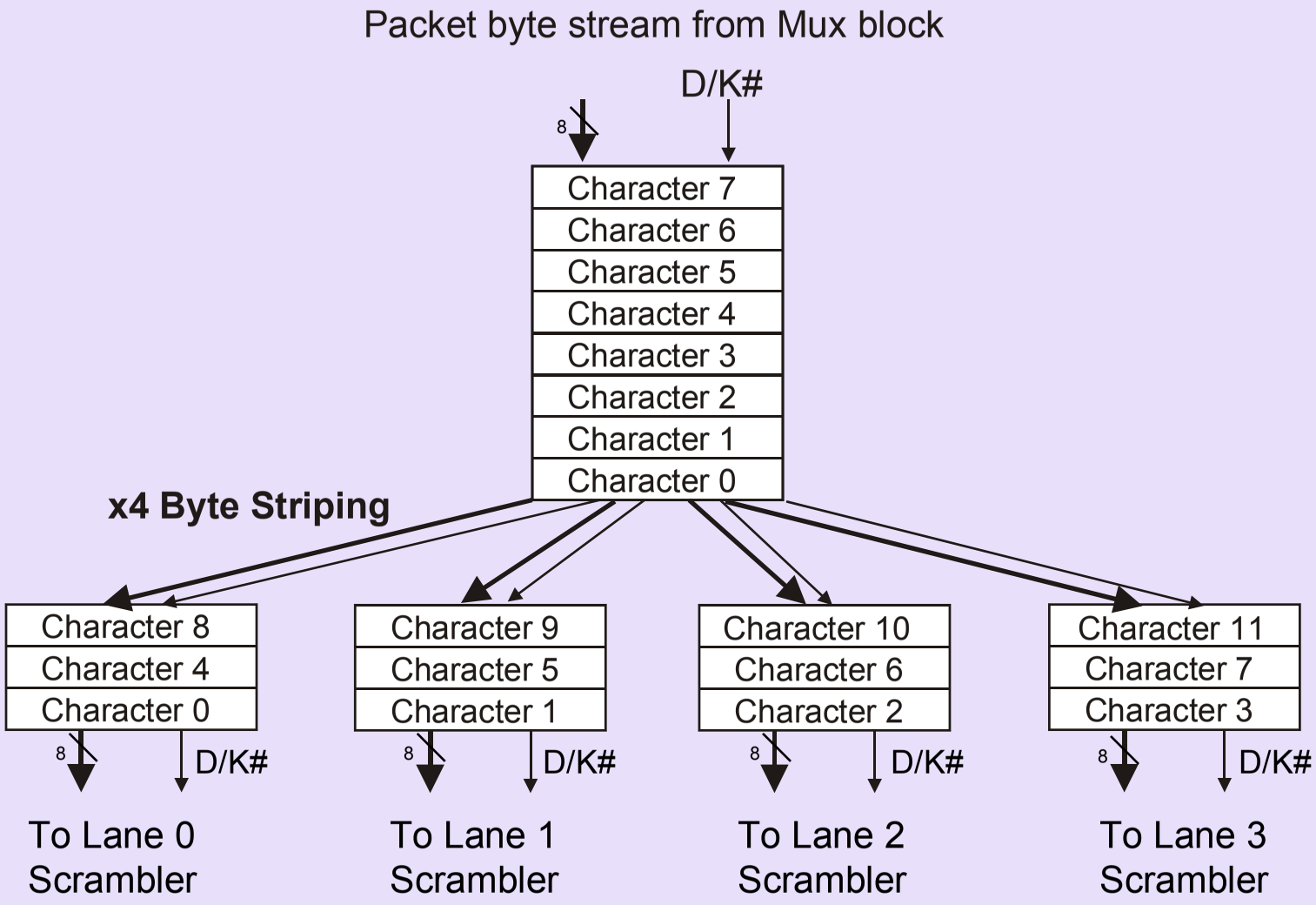
- **Electrical Functions**
  - ✓ Transmitter/receiver
  - ✓ Clocks/PLLs
  - ✓ Clock/data recovery

PHY layer upgrades do NOT impact upper layers

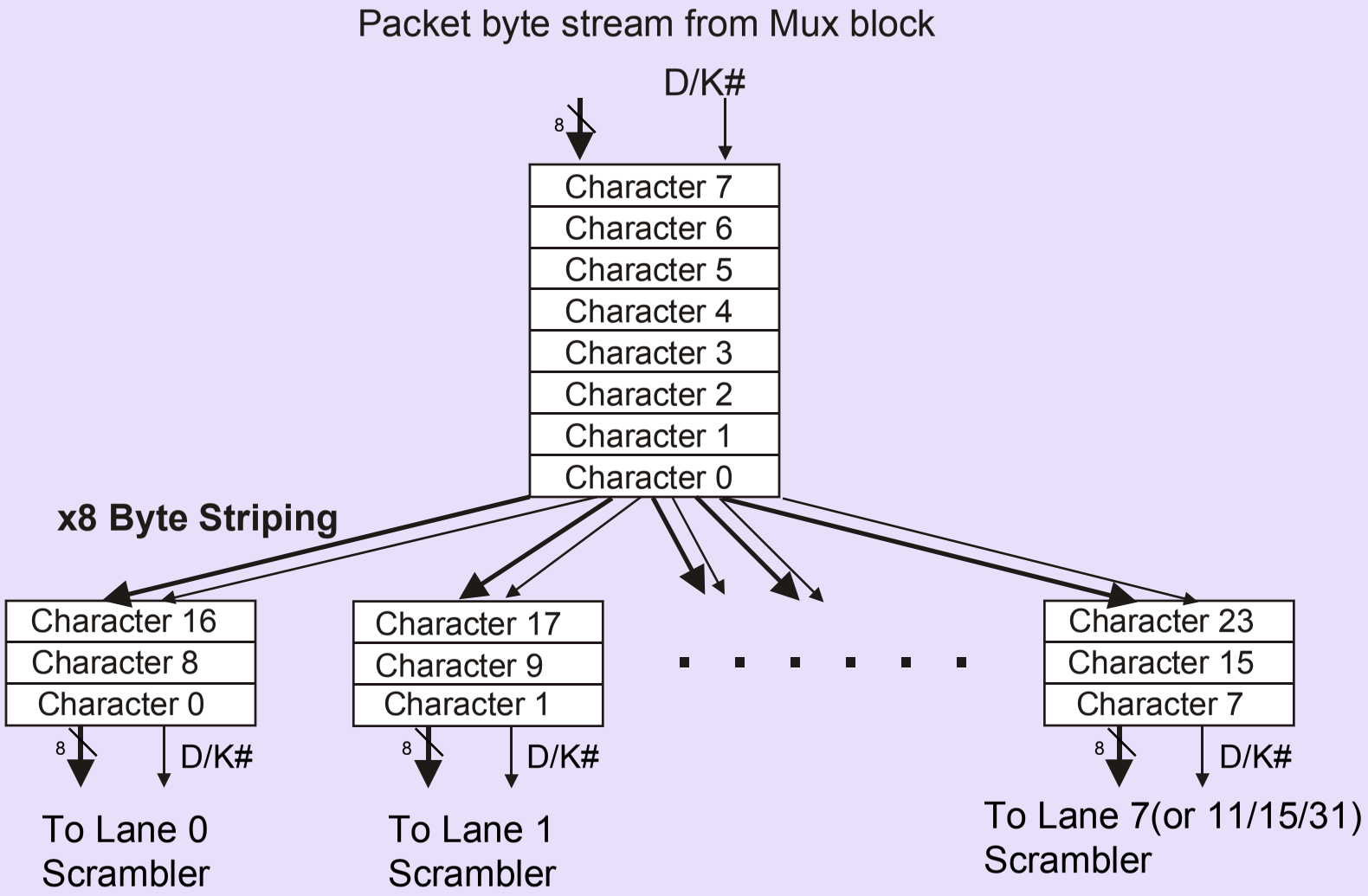
Logical Physical Layer



x4 Byte Striping



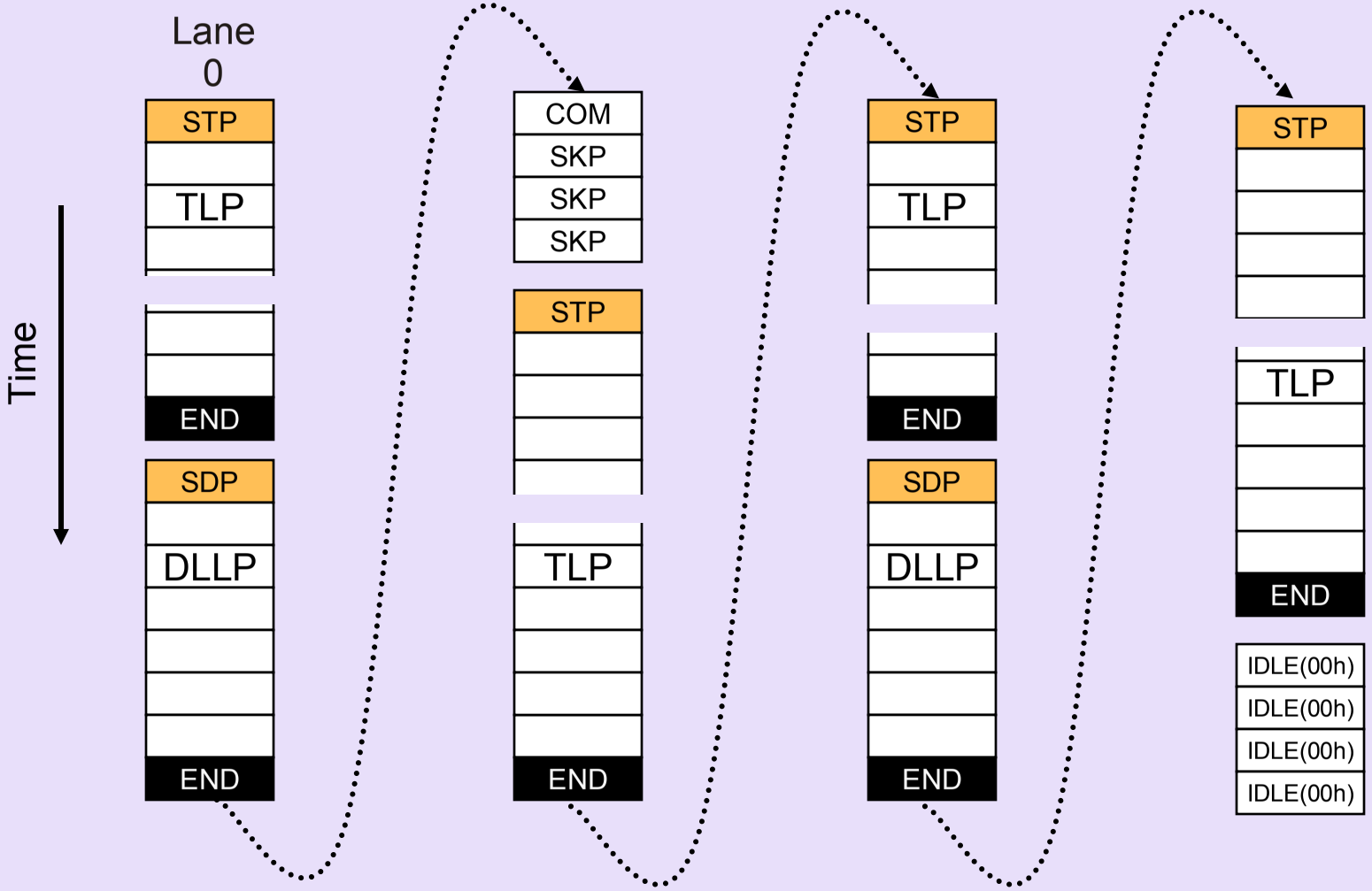
x8 Byte Striping



General Packet Format Rules

- TLP packets start with STP
- DLLP packets start with SDP and are 8 characters long (6 characters + SDP +END)
- All packets terminate with END or EDB
- When no packets are to be transferred, the logical Idle (00h) characters are transmitted
- Packets are multiples of 4 characters
- STP and SDP characters must be placed on Lane 0 when starting transmission from logical Idle

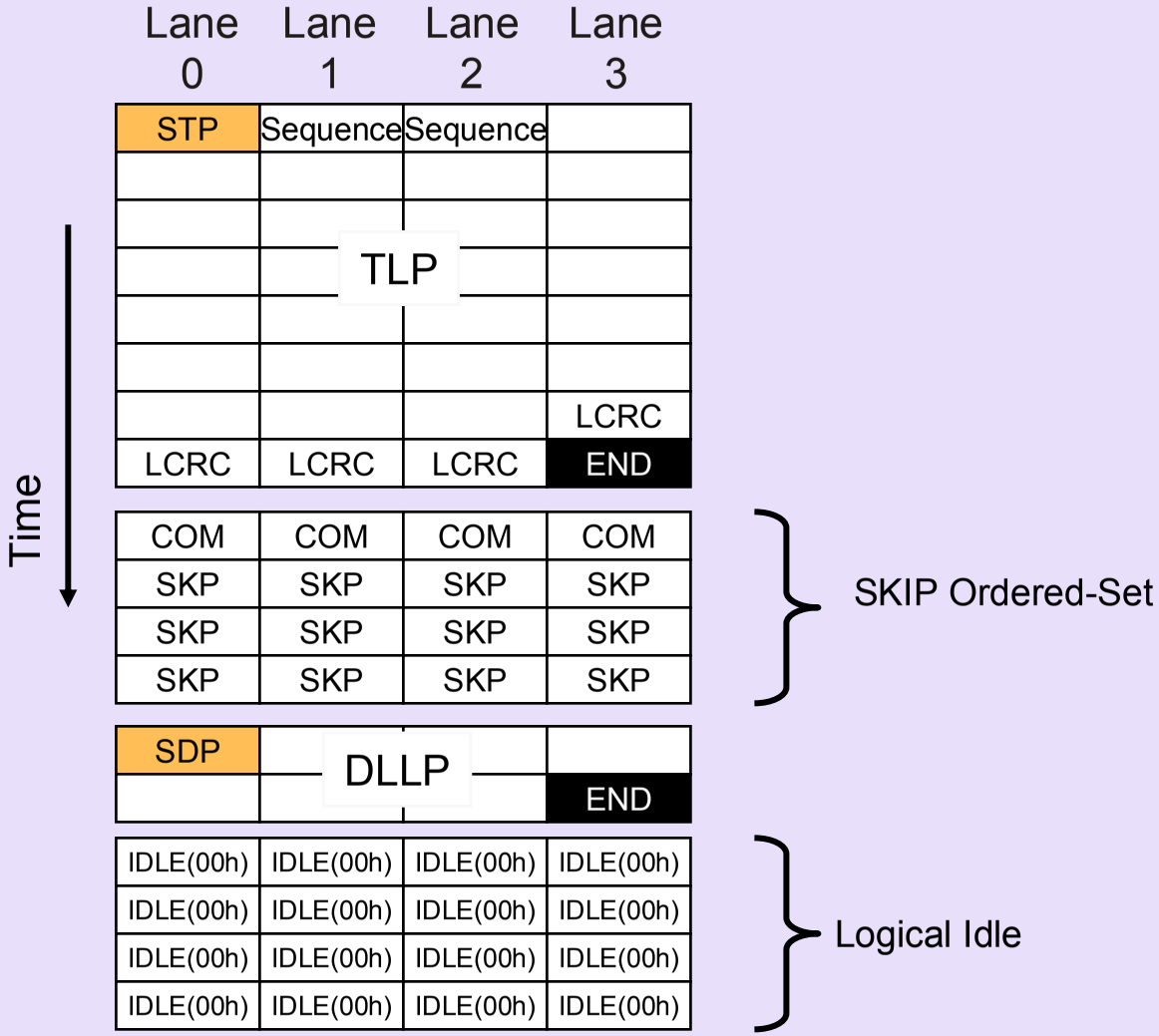
x1 Packet Format



x4 Packet Format Rules

- STP/SDP characters transmitted in Lane 0
- END/EDB characters transmitted in Lane 3
- When SKIP ordered-set is transmitted, it is transmitted on all 4 Lanes
- When logical Idle data is transmitted, it is transmitted on all Lanes

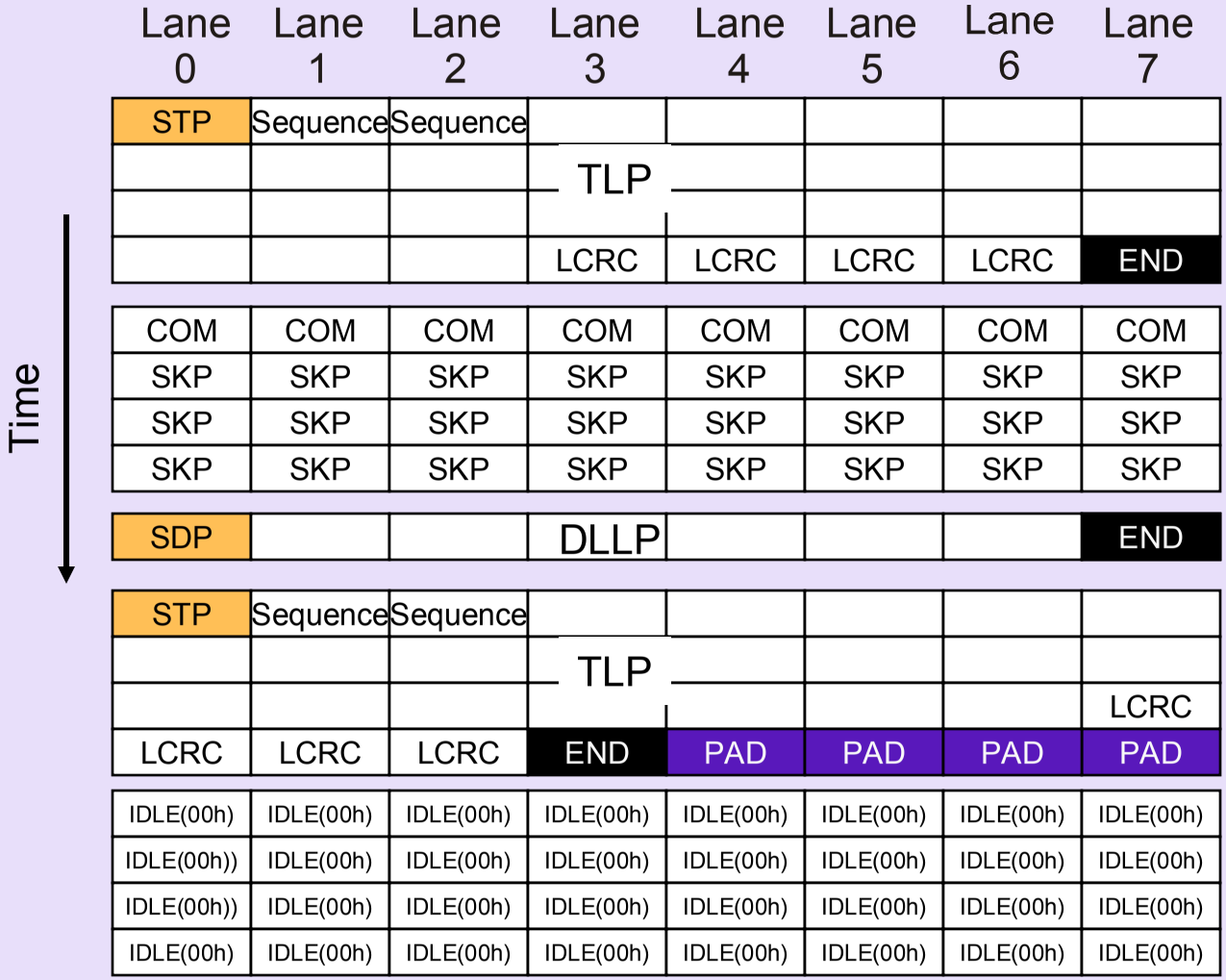
x4 Packet Format



x8, x12, x16, x32 Packet Format Rules

- STP/SDP characters are transmitted in Lane 0
  - ✓ They can also be transmitted in Lane 4*N (N>0) for consecutive transmissions of TLPs and DLLPs
- END/EDB characters may be transmitted in Lanes 3, 7, 11, 15, 19, 23, 27 or 31
- When packet does not end on last Lane and no other transmission, then PAD characters are added till last Lane to align Link for next transmission
- When SKIP ordered-set is transmitted, it is transmitted on all Lanes
- When logical Idle data is transmitted, it is transmitted on all Lanes

x8 Packet Format

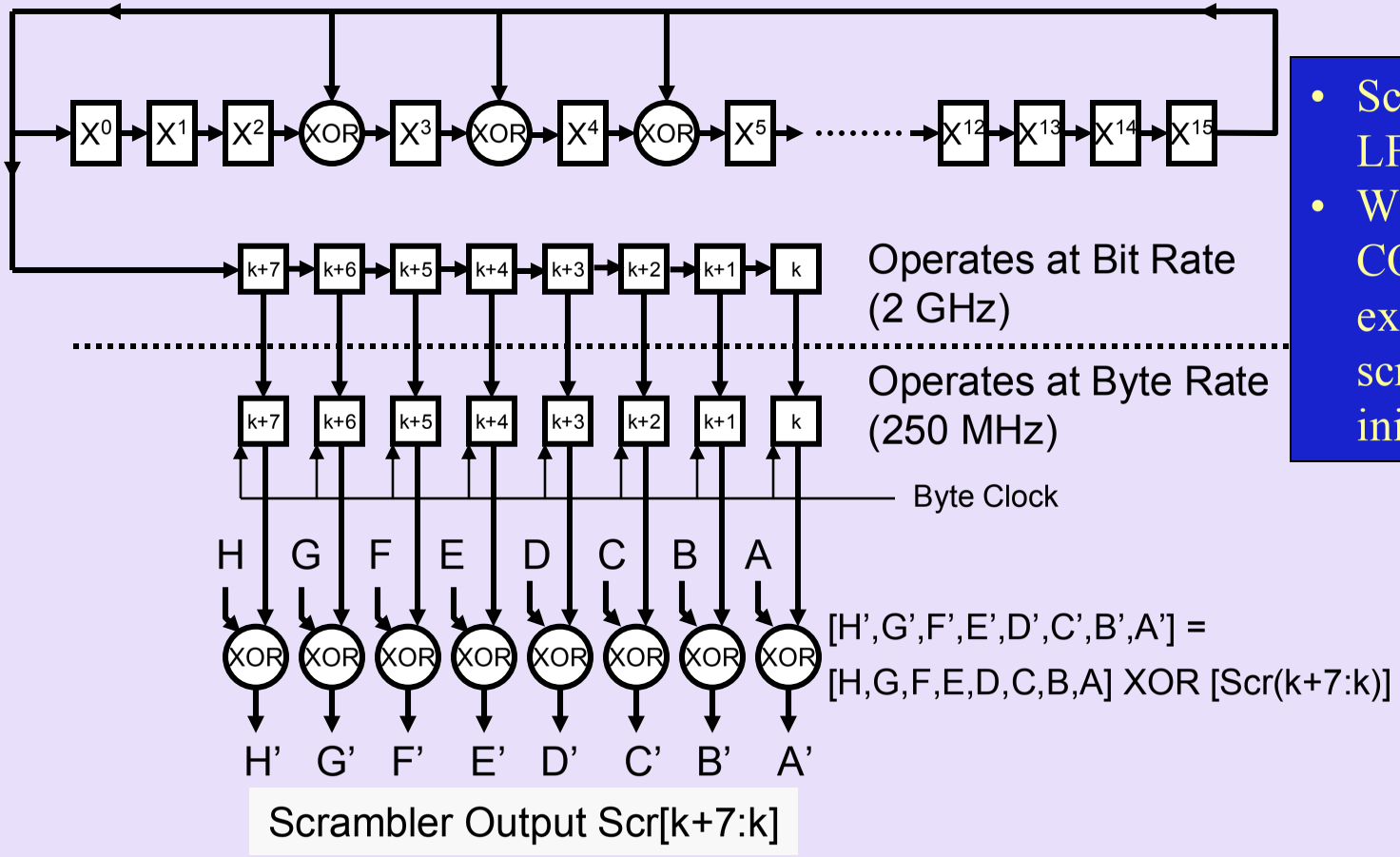


PAD characters are transmitted to maintain packet framing alignment

Scrambler Operation

- Scrambling polynomial:

$$G(x) = X^{16}+X^5+X^4+X^3+1$$

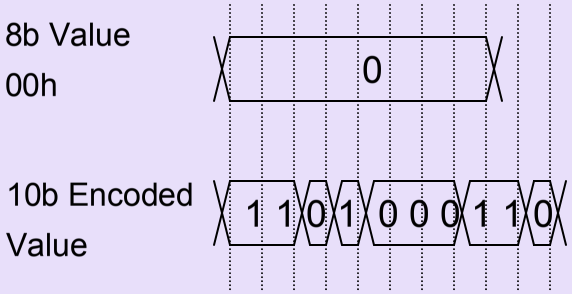


- Scrambler uses LFSR
- Whenever COM character exits scrambler, scramble is initialized

8b/10b Tutorial

- Standard encoding method invented by IBM and used in the following standards also:
  - ✓ Ethernet, Fibre Channel, ServerNet, FICON, IBA
- Reasons for encoding:
  - ✓ Creates sufficient transition density to limit “Run Length”
  - ✓ Balance number of 1’s and 0’s to maintain “DC Balance”
  - ✓ Ability to encode for special control characters (K)
  - ✓ Facilitate detection of most transmission errors
- Transmission performance degraded by 20%

8b/10b Encoding Example



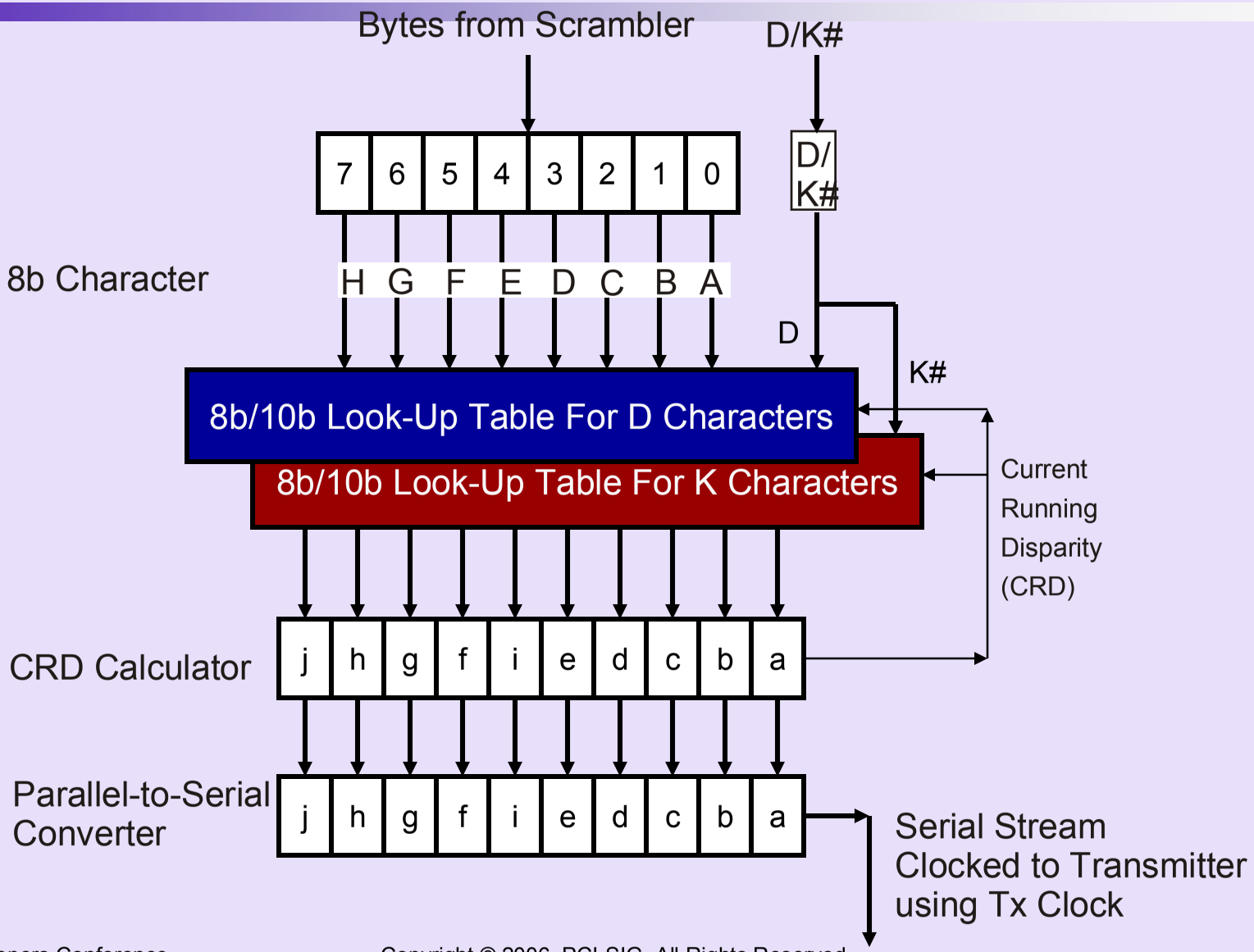
Properties of 10b Symbols

- Generally equal number of 1's and 0's over 2 consecutive symbols
- Maximum of five continuous strings of 1's or 0's
- Each 10b symbol has:
  - ✓ Four 0's and six 1's or,
  - ✓ Six 0's and four 1's or
  - ✓ Five 0's and five 1's
- 6 bit sub-block of 10b symbol contain no more than four 1's or four 0's
- 4 bit sub-block of 10b symbol contains no more than three 1's or three 0's
- Any other symbol having other than the above five properties is considered invalid

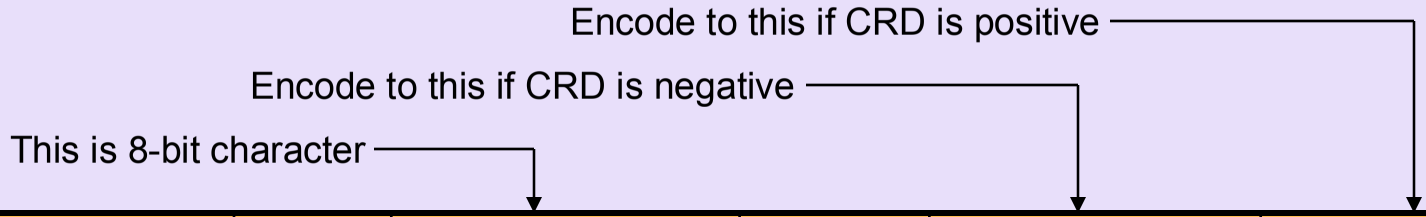
"Disparity" in 8b/10b Encoding

- Disparity is the difference between the number of 1's and 0's in a 10b symbol
- Symbols with:
  - ✓ Four 0's and six 1's has + disparity (CRD+)
  - ✓ Six 0's and four 1's has - disparity (CRD-)
  - ✓ Five 0's and five 1's has neutral disparity
- 8b character encode to 2 possible 10b symbols based on "Current Running Disparity" (CRD). Either:
  - ✓ One encoding contains four 0's and six 1's while second encoding contains six 0's and four 1's or
  - ✓ Both encodings contain five 0's and five 1's (neutral disparity)

8b/10b Encoder and Parallel-to-Serial Converter



Some Example 8b/10b Encoding



D or K Character	Hex Byte	Binary Bits HGF EDCBA	Byte Name	CRD – abcdei fghj	CRD + abcdei fghj
Data (D)	6A	011 01010	D10.3	010101 1100	010101 0011
Data (D)	1B	000 11011	D27.0	110110 0100	001001 1011
Data (D)	F7	111 10111	D23.7	111010 0001	000101 1110
Control (K)	F7	111 10111	K23.7	111010 1000	000101 0111
Control (K)	BC	101 11100	K28.5	001111 1010	110000 0101

If character encode yields neutral disparity, then CRD remains unchanged, else it flips

Example 8b/10b Transmission

Use these two characters in the example below:

D/K#	Hex Byte	Binary Bits HGF EDCBA	Byte Name	CRD – abcdei fghj	CRD + abcdei fghj
Control (K)	BC	101 11100	K28.5	001111 1010	110000 0101
Data (D)	6A	011 01010	D10.3	010101 1100	010101 0011

Example Transmission

	CRD	Character	CRD	Character	CRD	Character	CRD
Character to be transmitted		K28.5 (BCh)		K28.5 (BCh)		D10.3 (6Ah)	
Bit stream transmitted	-	Yields 001111 1010 Symbol disparity is +	+	Yields 110000 0101 Symbol disparity is -	-	Yields 010101 1100 Symbol disparity is neutral	-

Initialized value of CRD is don't care. Receiver can determine from incoming bit stream

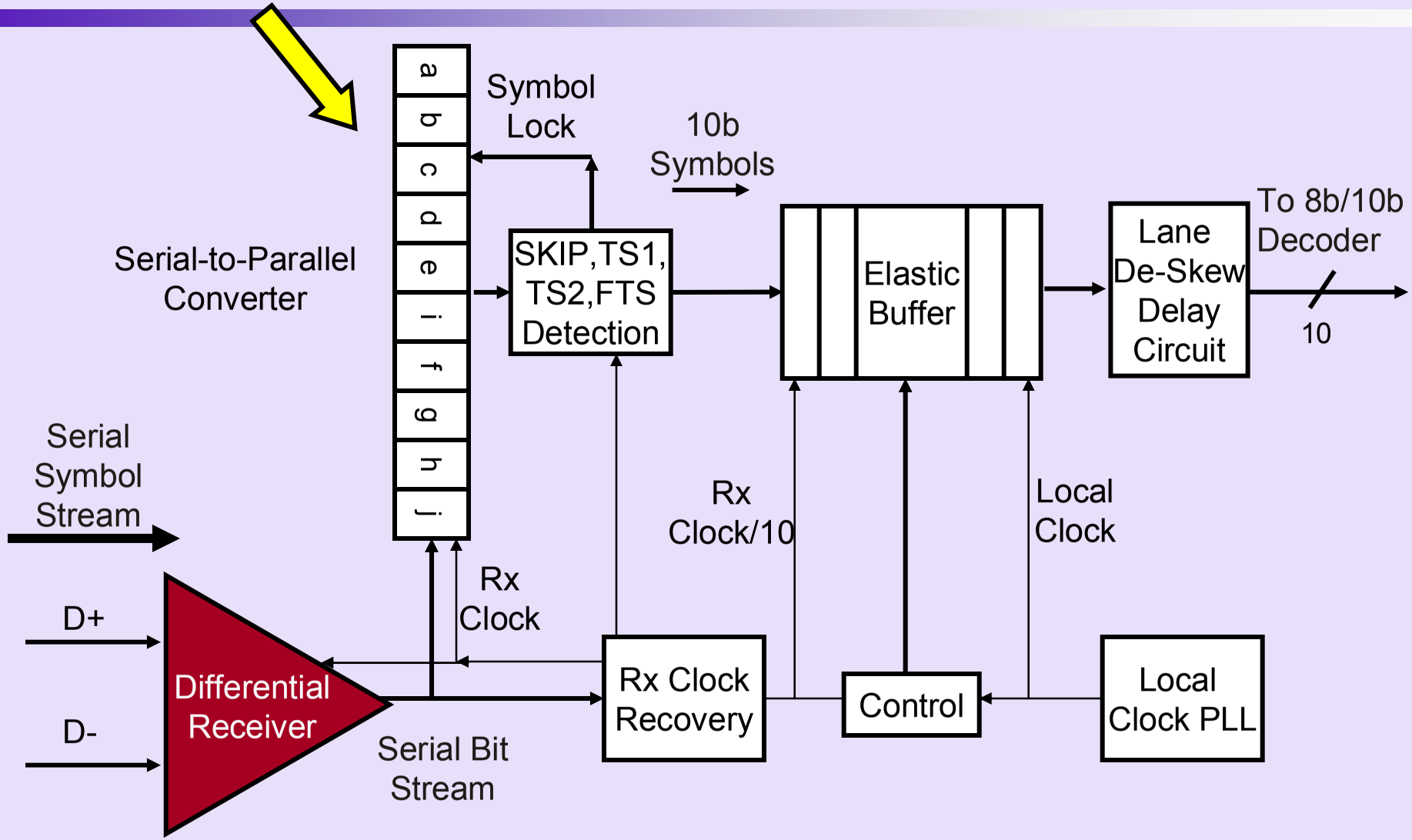
Control Characters and Their Symbol Encoding

Character	8b Name	10b (CRD-)	10b (CRD+)	Description
COM	K28.5 (BCh)	001111 1010	110000 0101	Comma used as a character boundary alignment symbol
PAD	K23.7 (F7h)	111010 1000	000101 0111	Packet Padding Symbol
SKP	K28.0 (1Ch)	001111 0100	110000 1011	Used in SKIP ordered-set
STP	K27.7 (FBh)	110110 1000	001001 0111	Start of TLP Symbol
SDP	K28.2 (5Ch)	001111 0101	110000 1010	Start of DLLP Symbol
END	K29.7 (FDh)	101110 1000	010001 0111	End of Good Packet Symbol
EDB	K30.7 (FEh)	011110 1000	100001 0111	End of Bad Packet Symbol Used by switch which detects bad packet
FTS	K28.1 (3Ch)	001111 1001	110000 0110	Used in ordered-set to exit L0s to L0 power state
IDL	K28.3 (7Ch)	001111 0011	110000 1100	Used in Electrical IDLE ordered-set

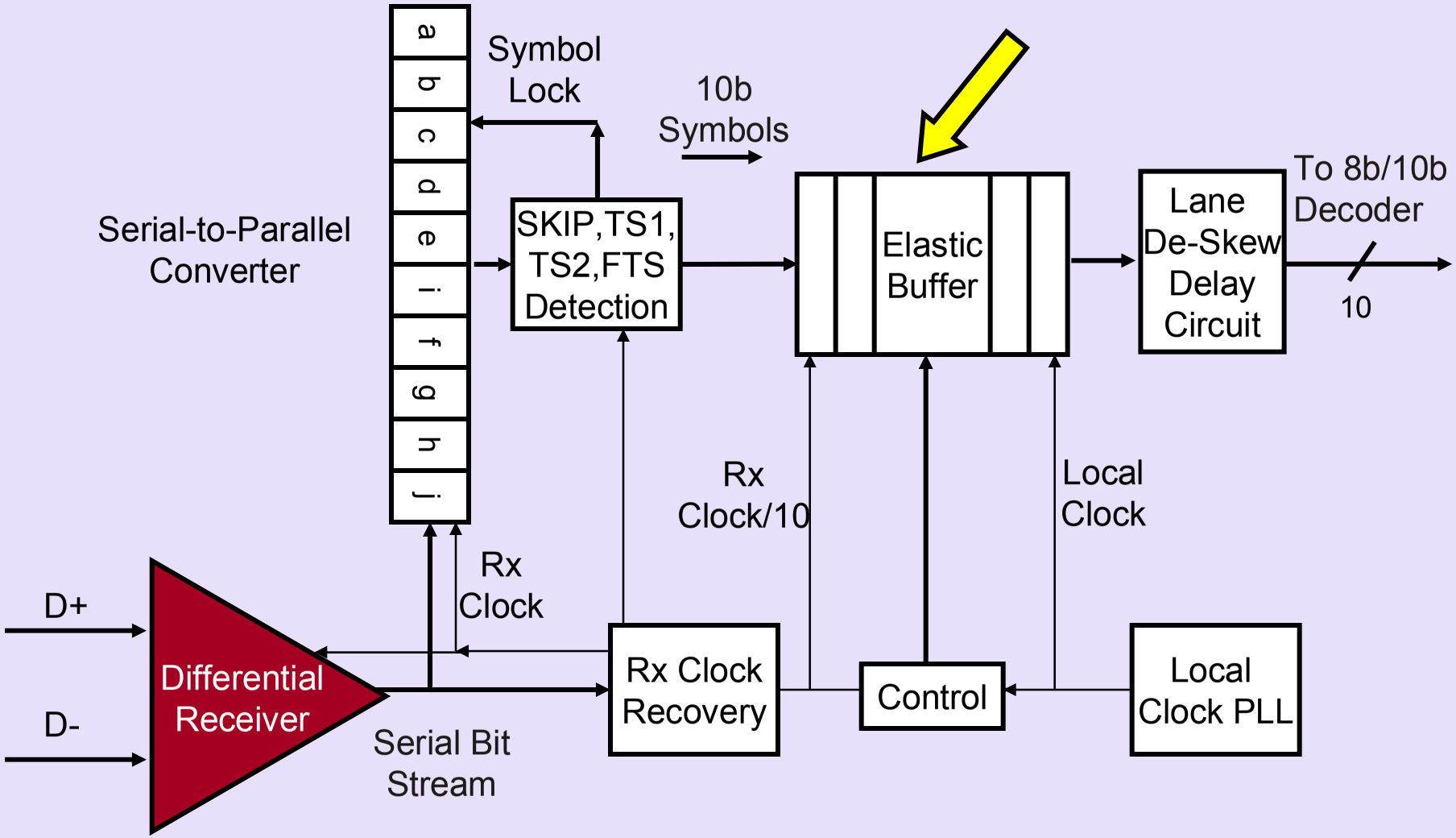
Comma Character Description

- COM character is the first symbol in a SKIP, TS1, TS2, Electrical IDLE, FTS ordered-set (discussed in next slide)
- 10b encoding of COM (K28.5) character contains 2 bits of one polarity followed by 5 bits of the opposite polarity (**001111 1**010 or **110000 0**101)
  - ✓ The only other symbol with this property is FTS
  - ✓ This property makes the COM symbol (also FTS) easily detectable

Serial-to-Parallel Converter



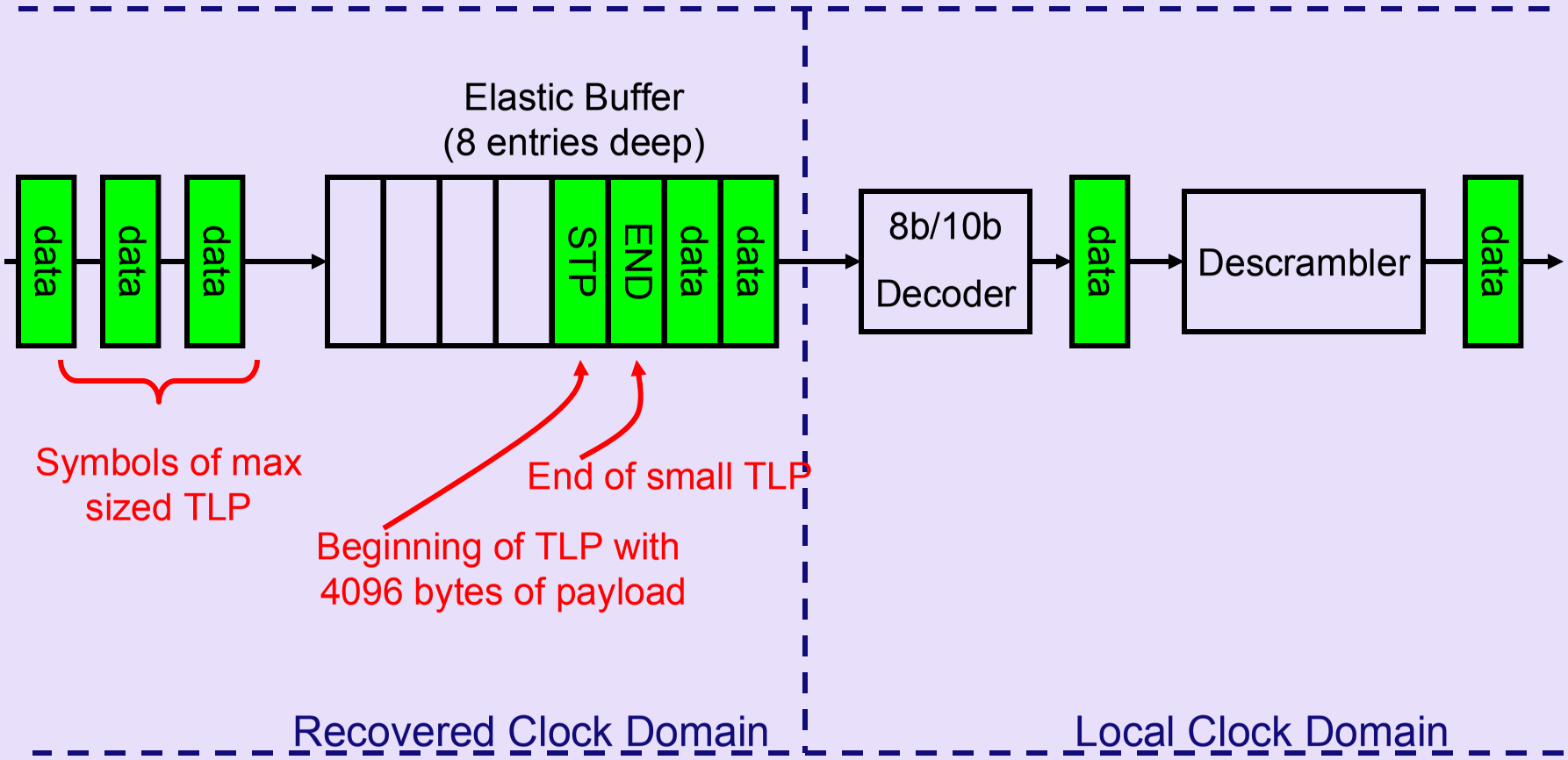
Elastic Buffer



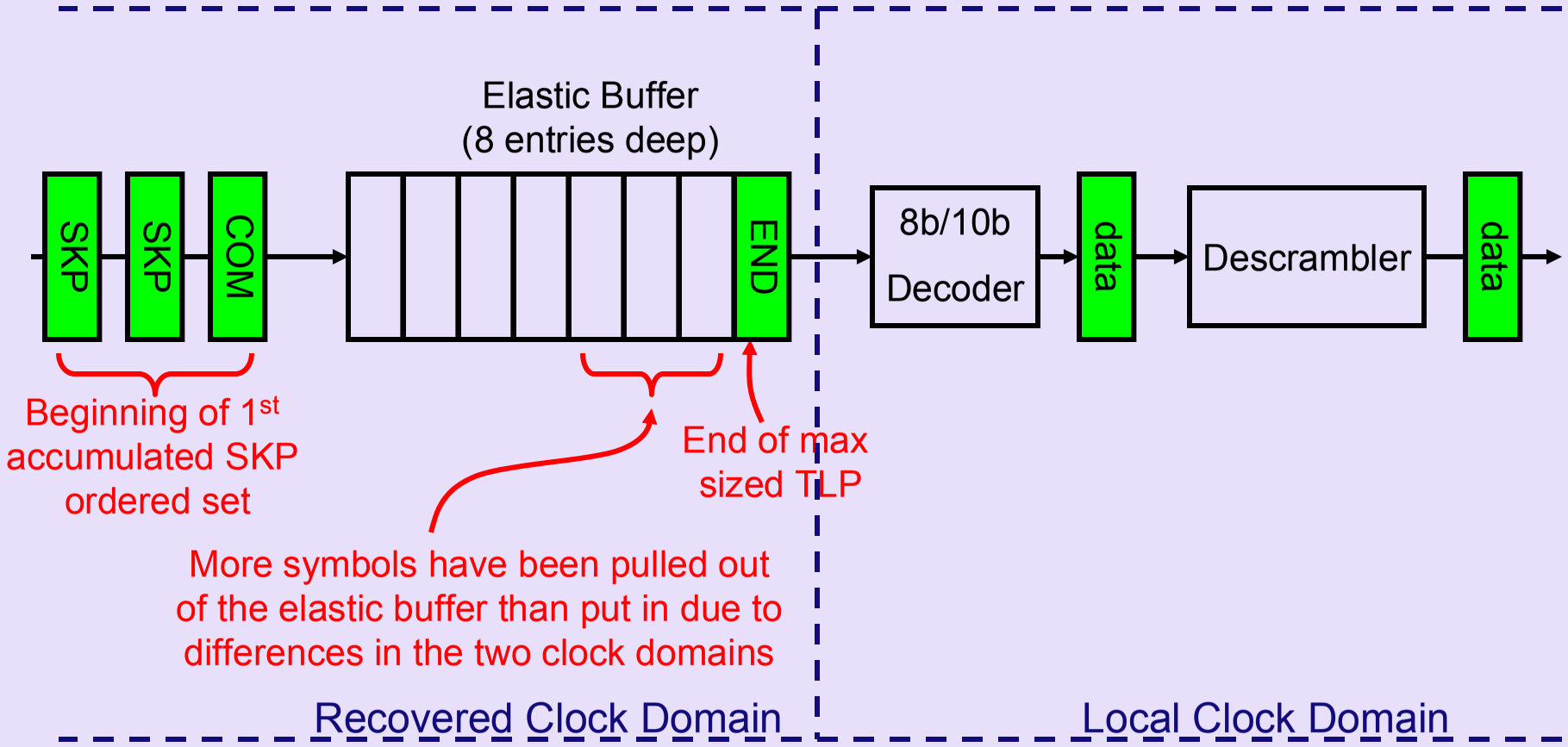
Clock Tolerance Compensation

- Skip ordered-set used to compensate for differences in frequency of bit rate a packet is transmitted with, and the receiver clock frequency
- Frequency difference no greater than 600 ppm
- Skip ordered-set scheduled to be inserted every 1180 to 1538 symbol times and transmitted on a packet boundary.

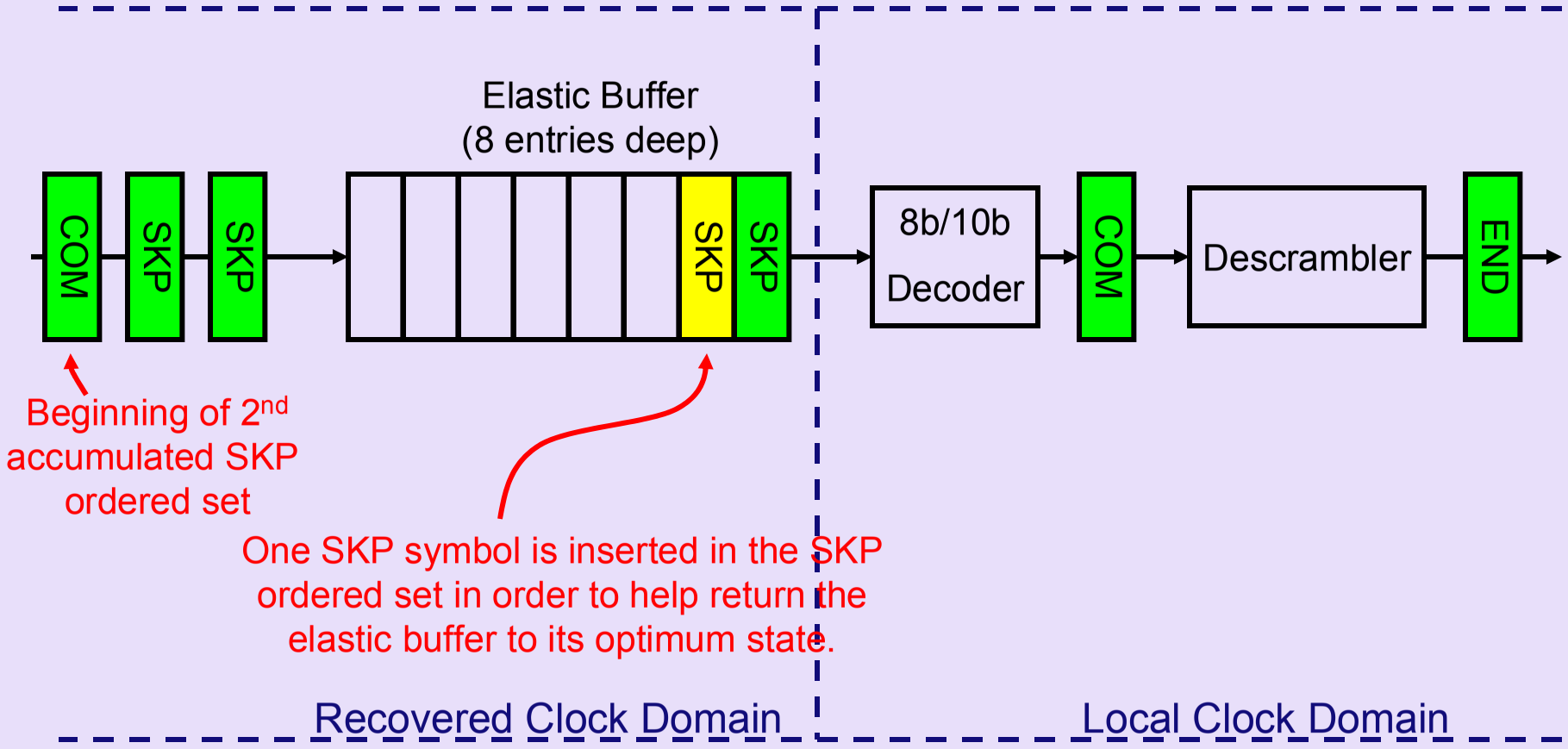
Elastic Buffer Option 1, Example 1: Local Clock Faster than Recovered Clock



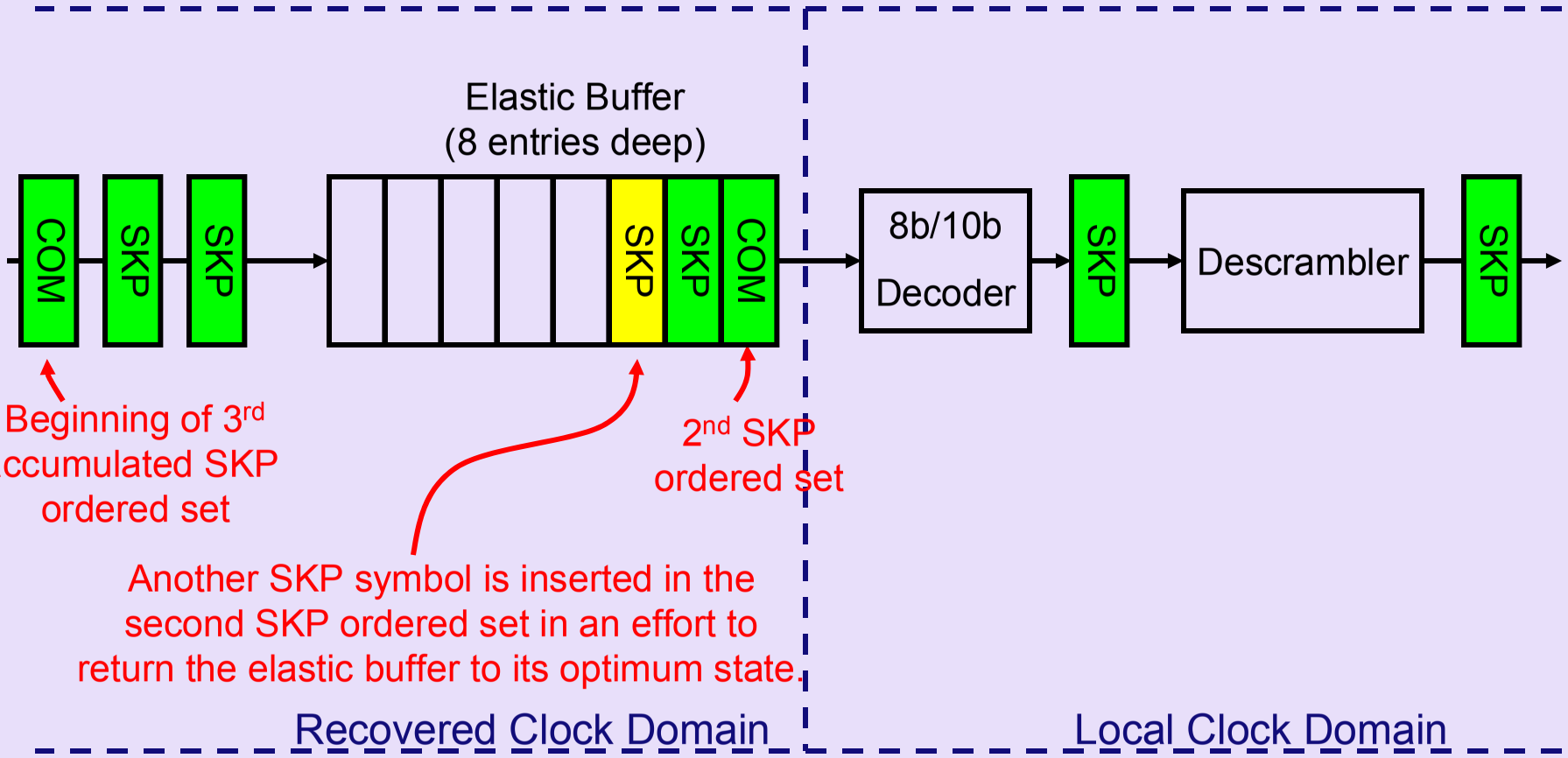
Elastic Buffer Option 1, Example 1: Local Clock Faster than Recovered Clock



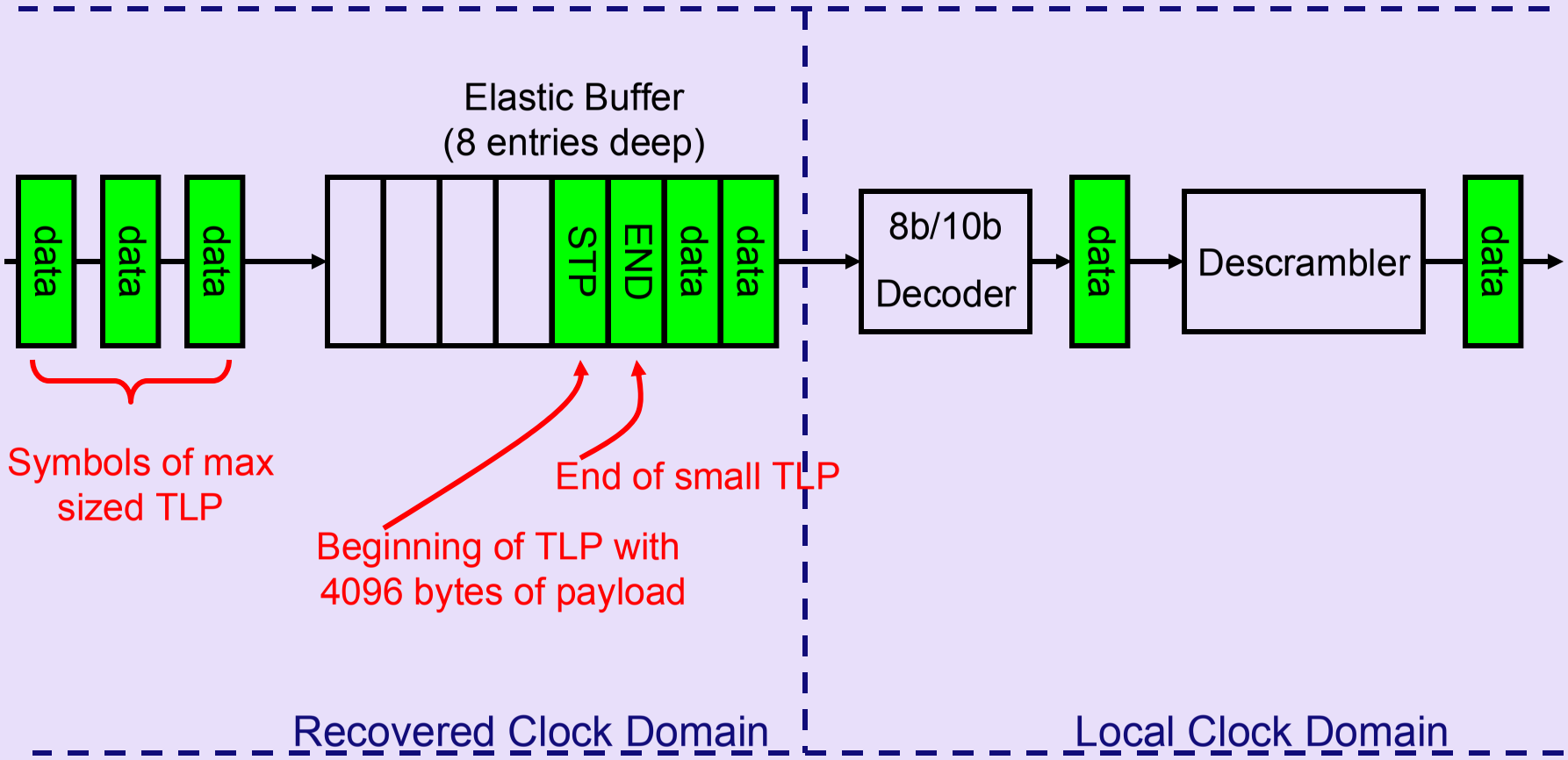
Elastic Buffer Option 1, Example 1: Local Clock Faster than Recovered Clock



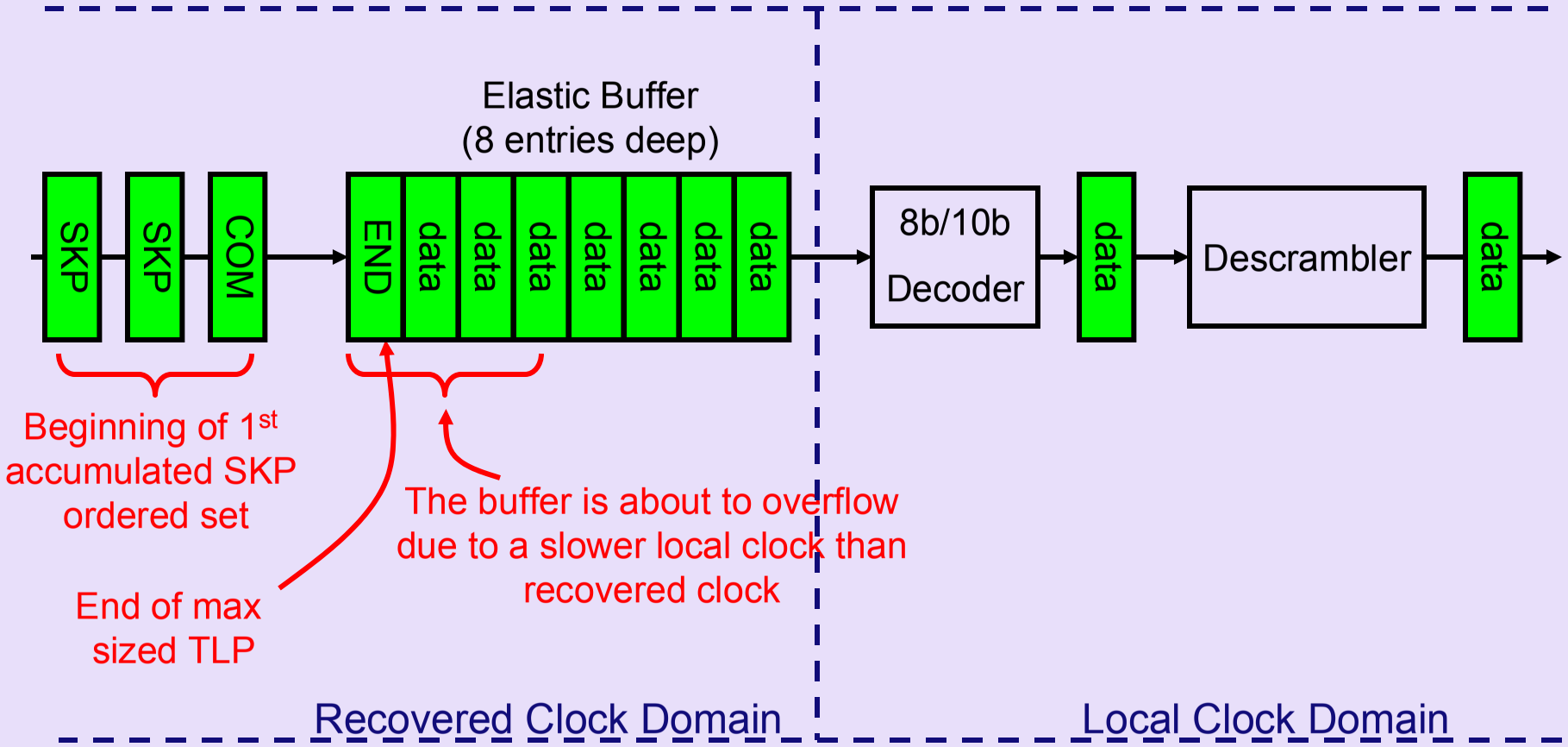
Elastic Buffer Option 1, Example 1: Local Clock Faster than Recovered Clock



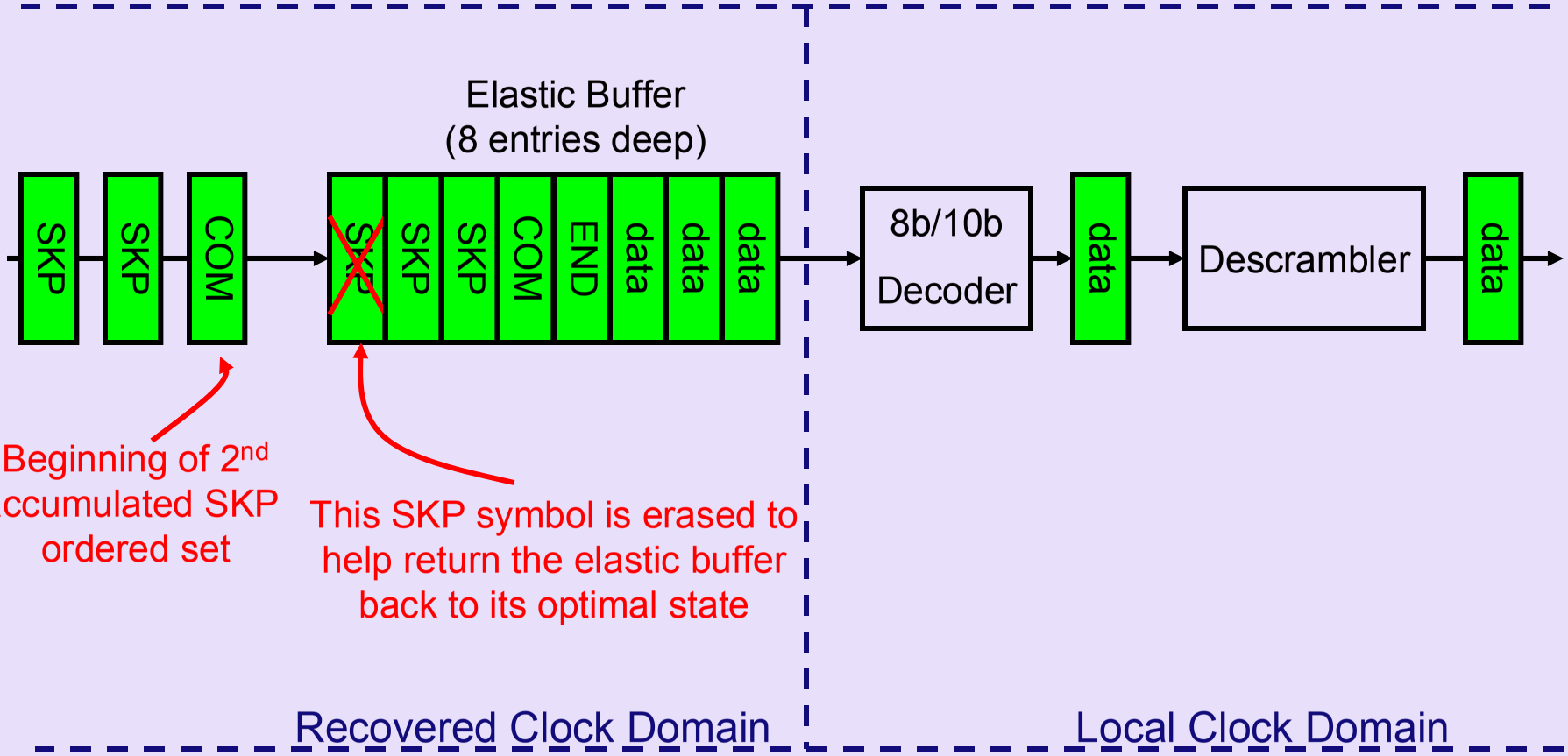
Elastic Buffer Option 1, Example 2: Local Clock Slower than Recovered Clock



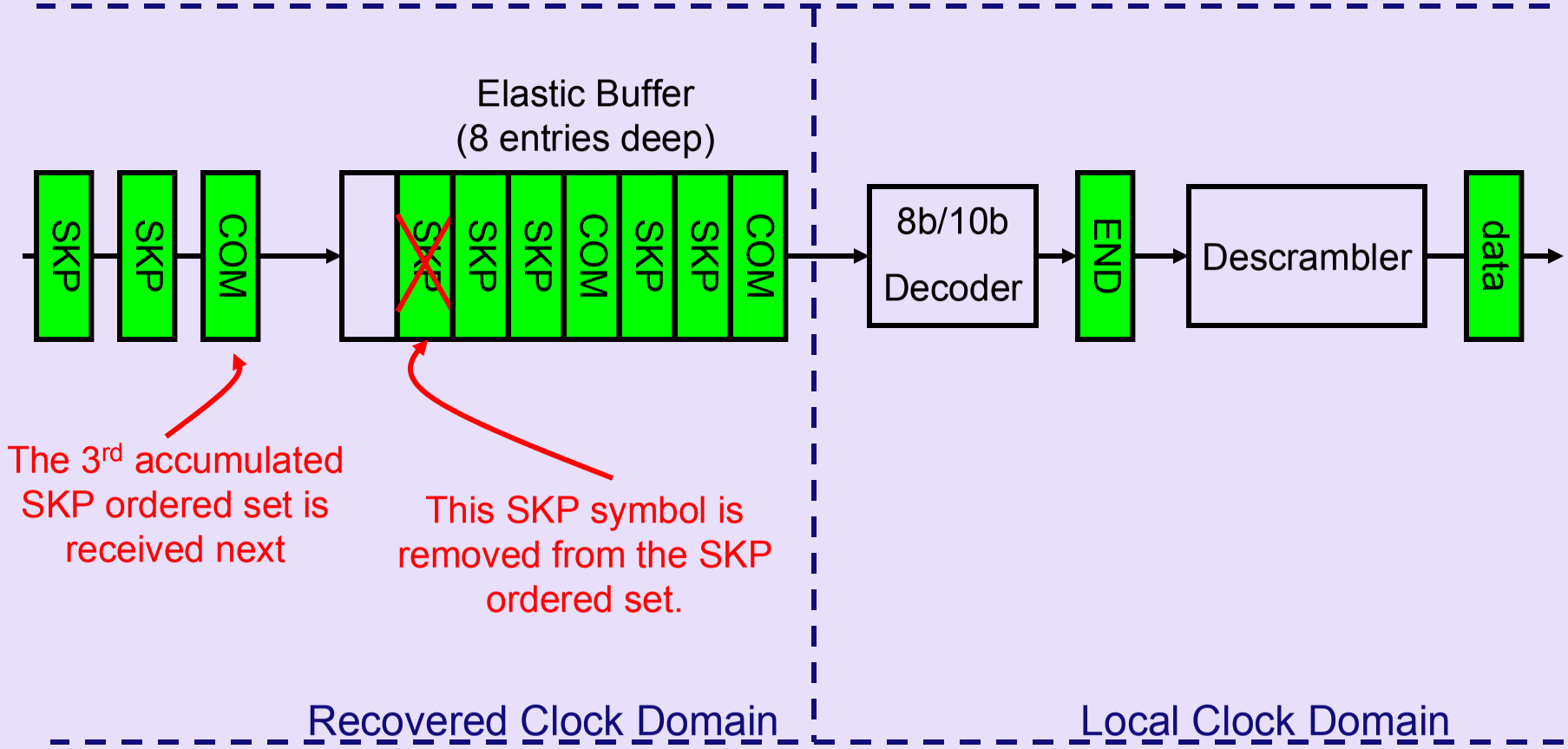
Elastic Buffer Option 1, Example 1: Local Clock Slower than Recovered Clock



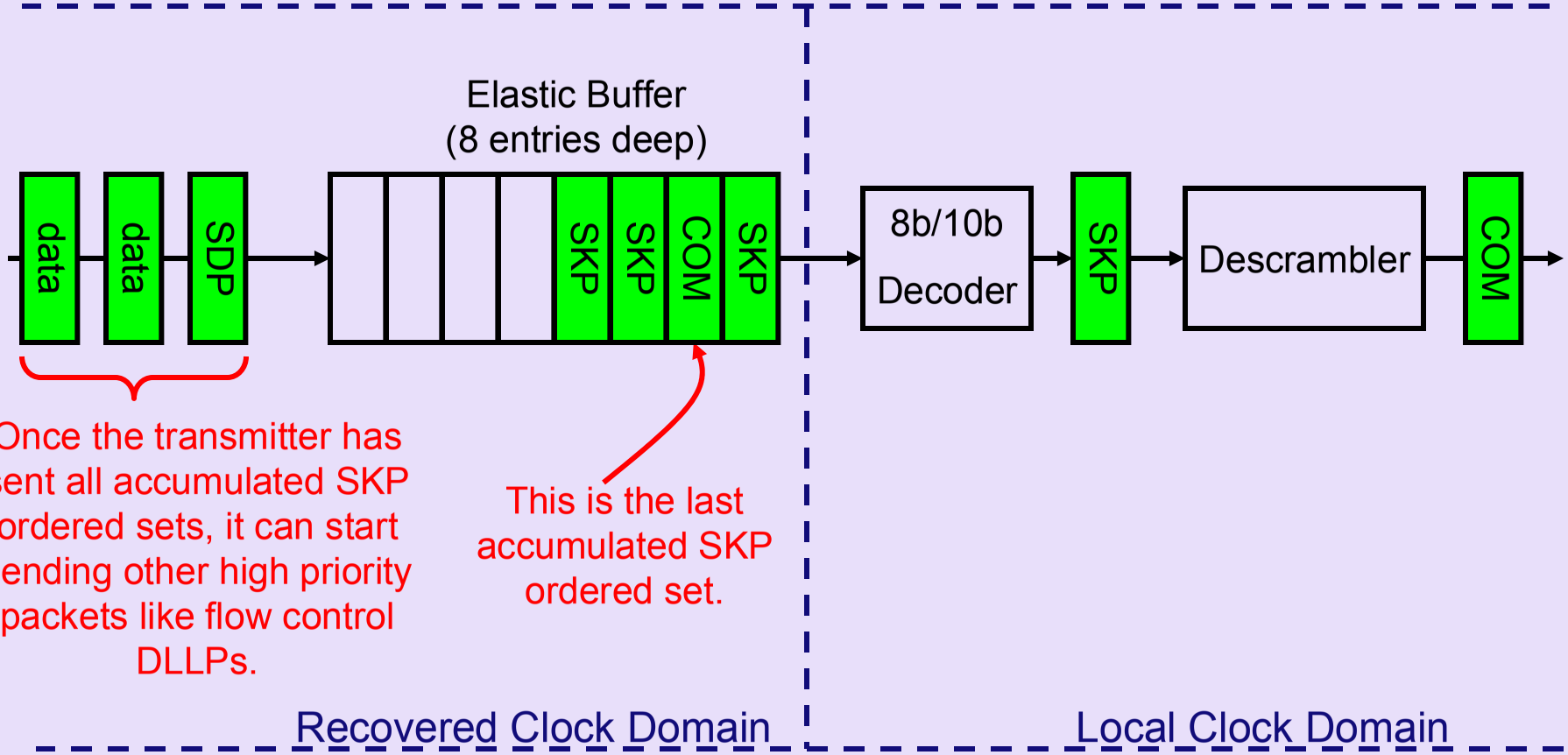
Elastic Buffer Option 1, Example 1: Local Clock Slower than Recovered Clock



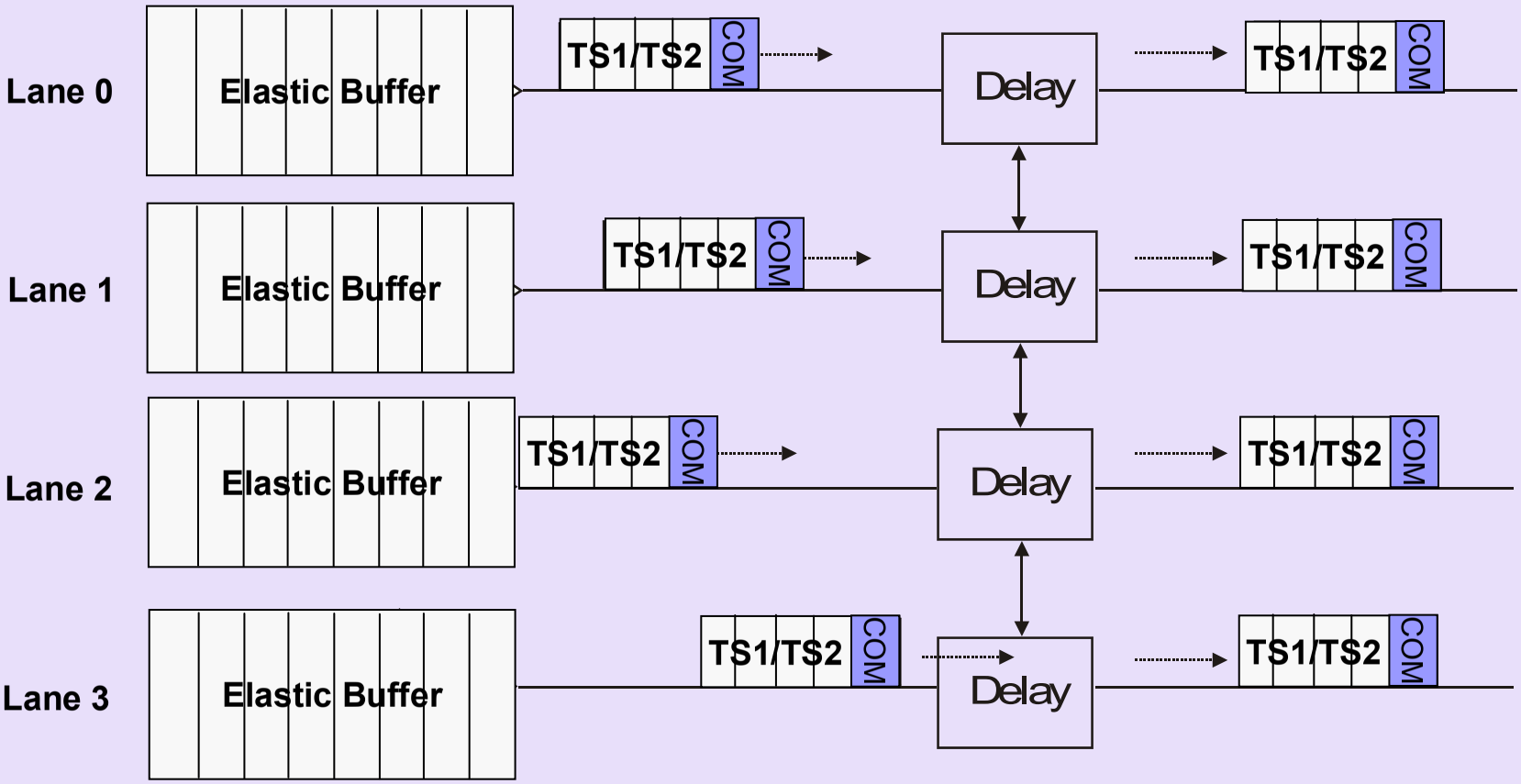
Elastic Buffer Option 1, Example 1: Local Clock Slower than Recovered Clock



Elastic Buffer Option 1, Example 1: Local Clock Slower than Recovered Clock



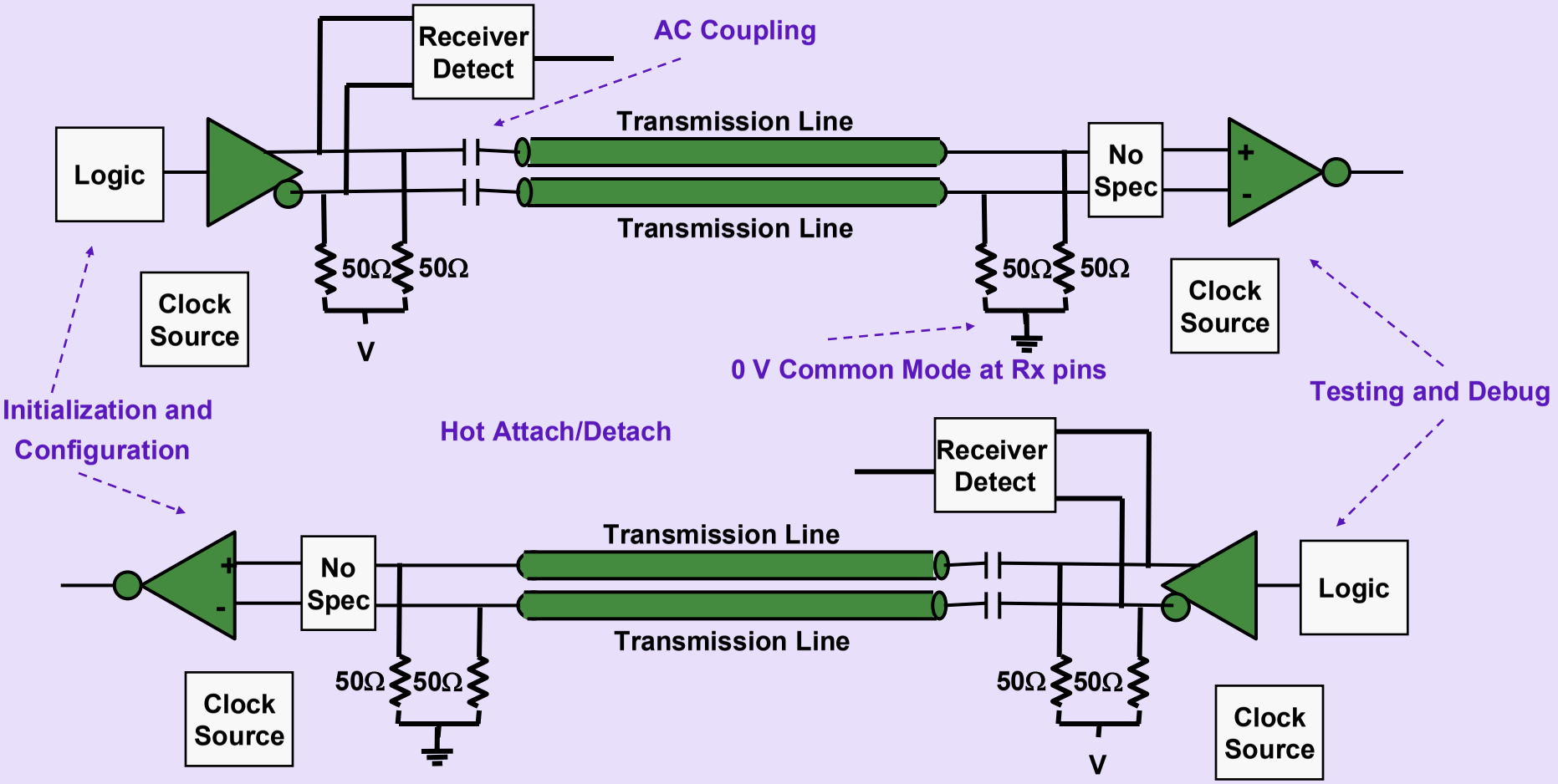
Receiver Link De-Skew



Agenda

- Logical Physical Layer
- Electrical Physical Layer
- Link Training and Initialization (LTSSM)
- Layout Considerations

Pictorial Summary

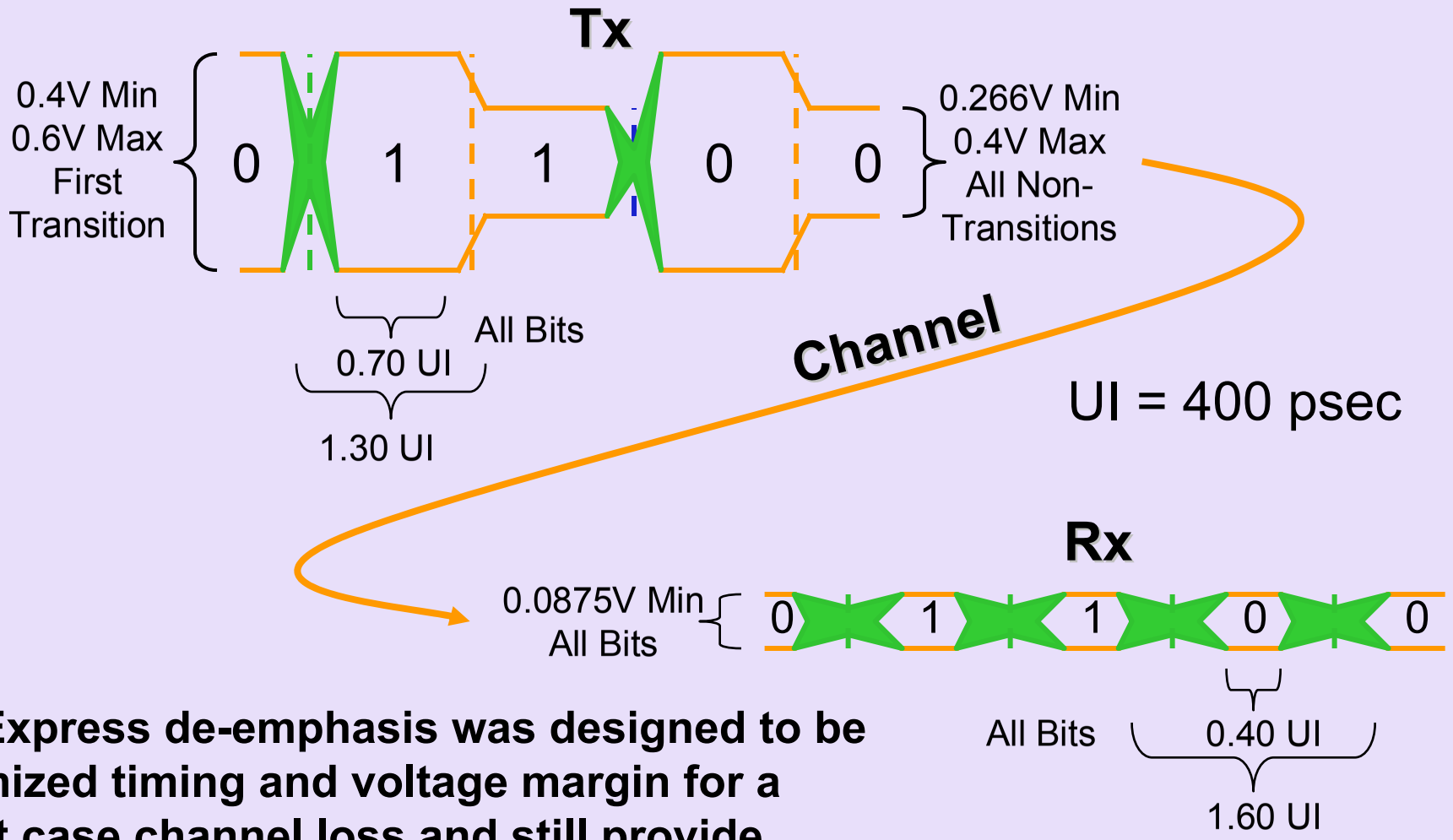


Emphasis on Low Power

ESD, EMI, Noise Suppression

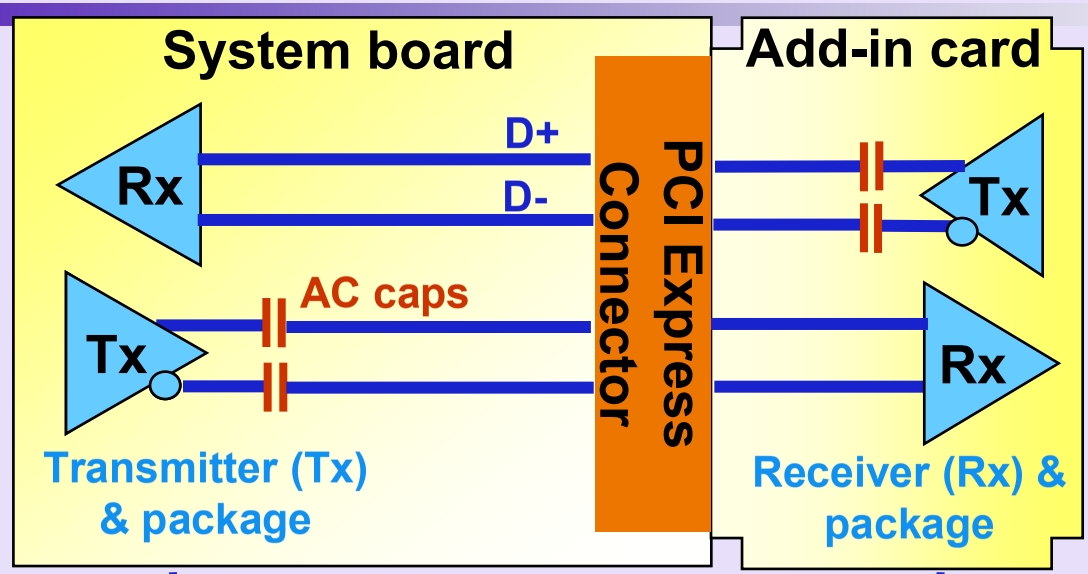
Error Behavior/Recovery

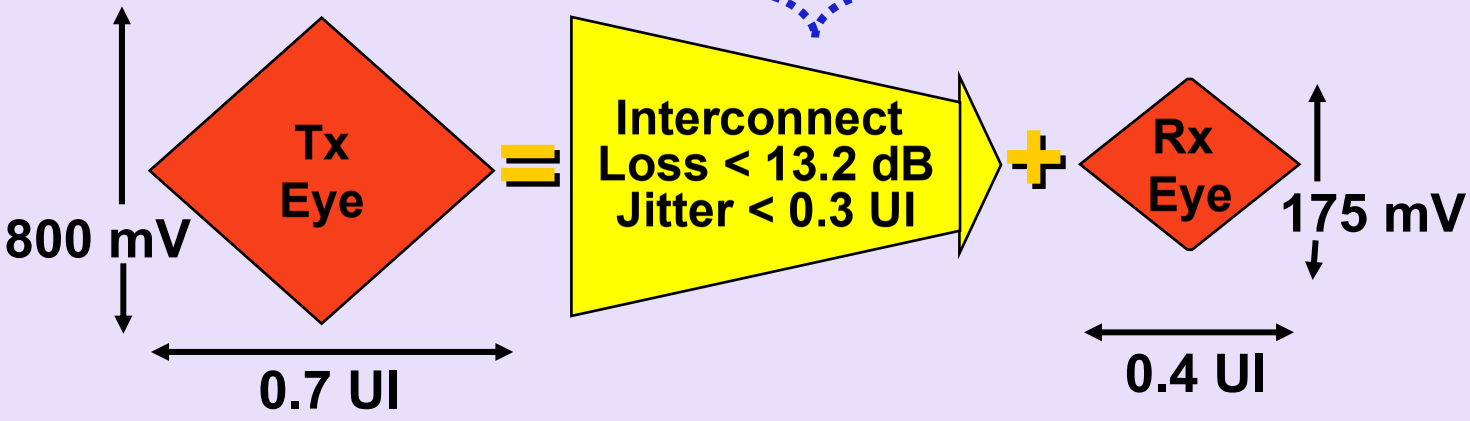
Electrical Specifications



PCI Express de-emphasis was designed to be optimized timing and voltage margin for a worst case channel loss and still provide adequate margin for the best case channel loss

System Budget



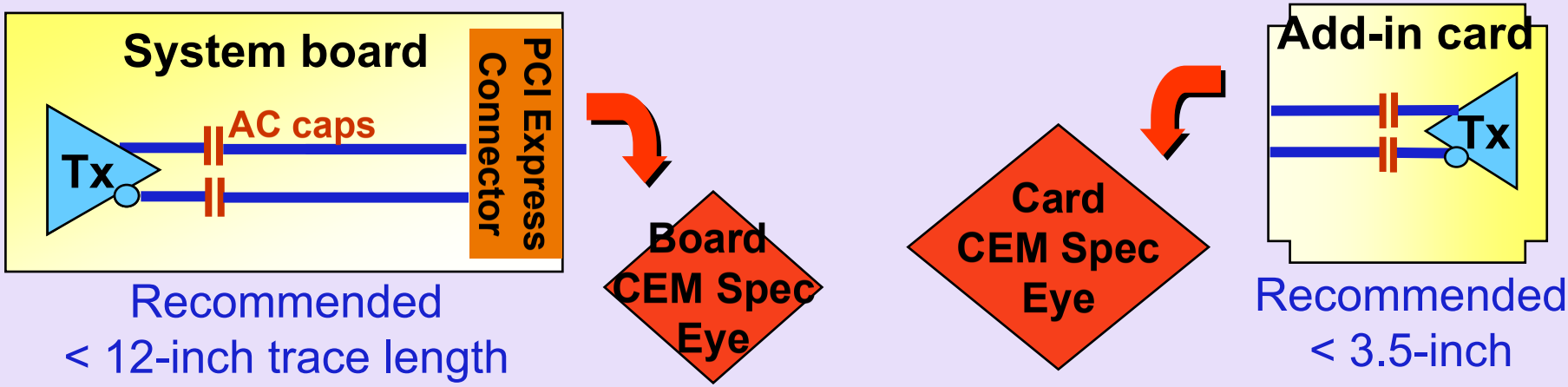


- ✓ Differential pairs
- ✓ AC coupled
- ✓ Lane-to-lane de-skew
- ✓ Polarity inversion
- ✓ On-chip equalization
- ✓ On-chip terminations

UI = Unit Interval as defined in the PCI Express Base 1.0a Specification

Card Electromechanical Interconnect Budget

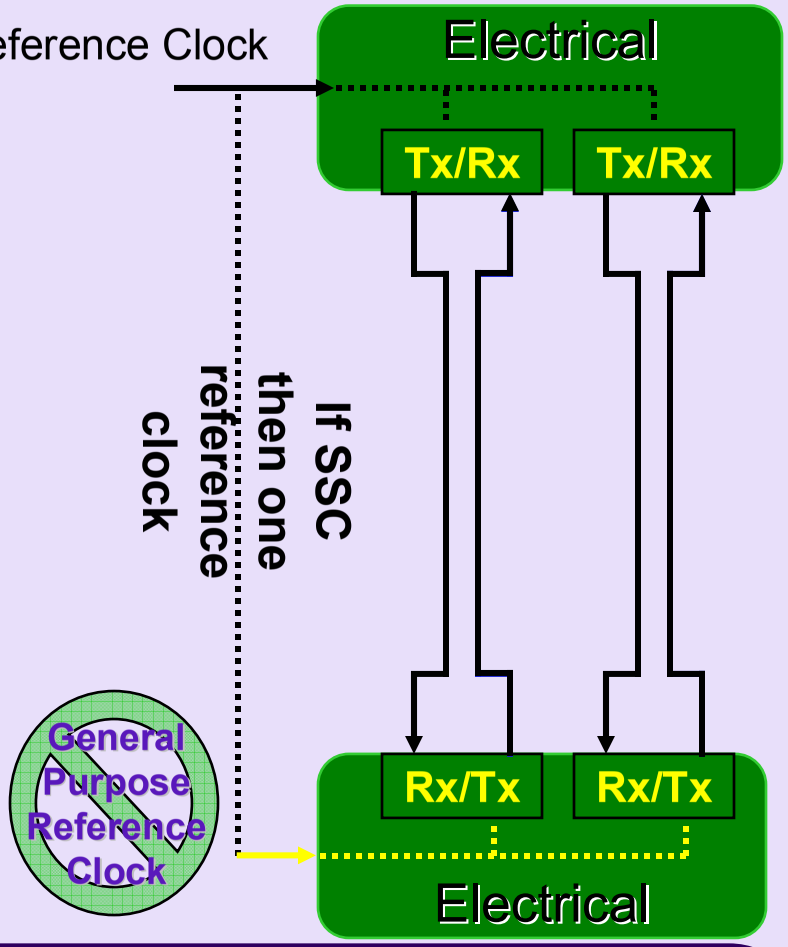
- **Card Electromechanical (CEM) 1.1 specification defines budget allocation**
  - ✓ *Loss* and *jitter* are key parameters
  - ✓ Target impedance not as critical
  - ✓ Maintain differential pair symmetry
  - ✓ Design tradeoffs: loss vs. trace length, etc.



Manage loss & jitter to meet budget

Clocking Options

- **All lanes within a port must transmit data using one frequency**
- **The ports at each end of a link may transmit data at slightly different frequencies**
  - ✓ Tolerance = ±300 ppm each

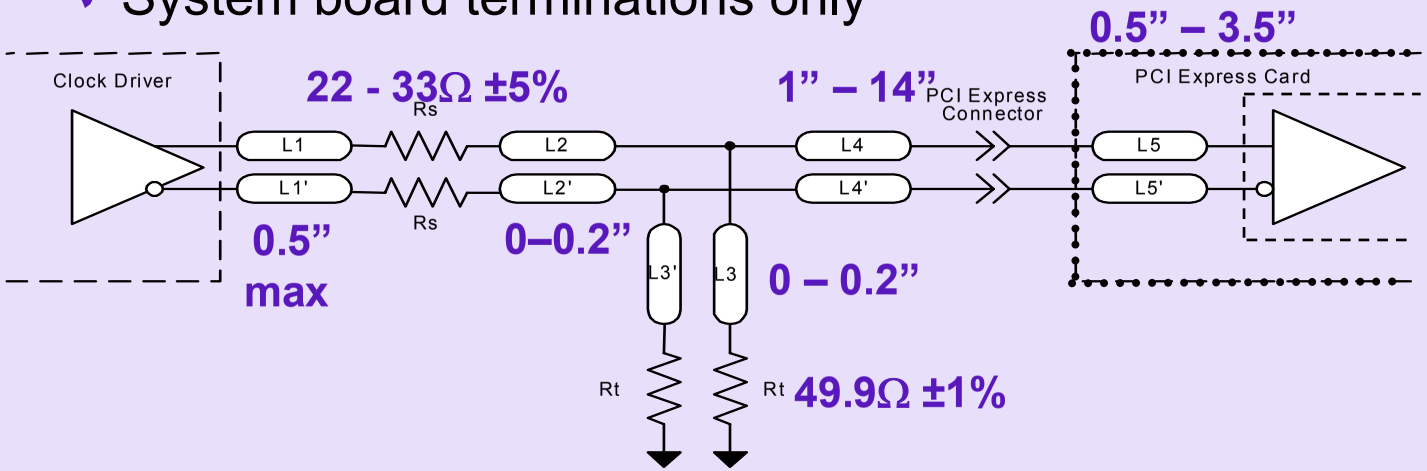


If SSC used to modulate data rate, then both ports should use same modulated clock source

Reference Clock Routing

- **Differential clock routing to each device and connector**
  - ✓ Use the same differential trace geometries
  - ✓ Length matching to different devices *NOT* required!

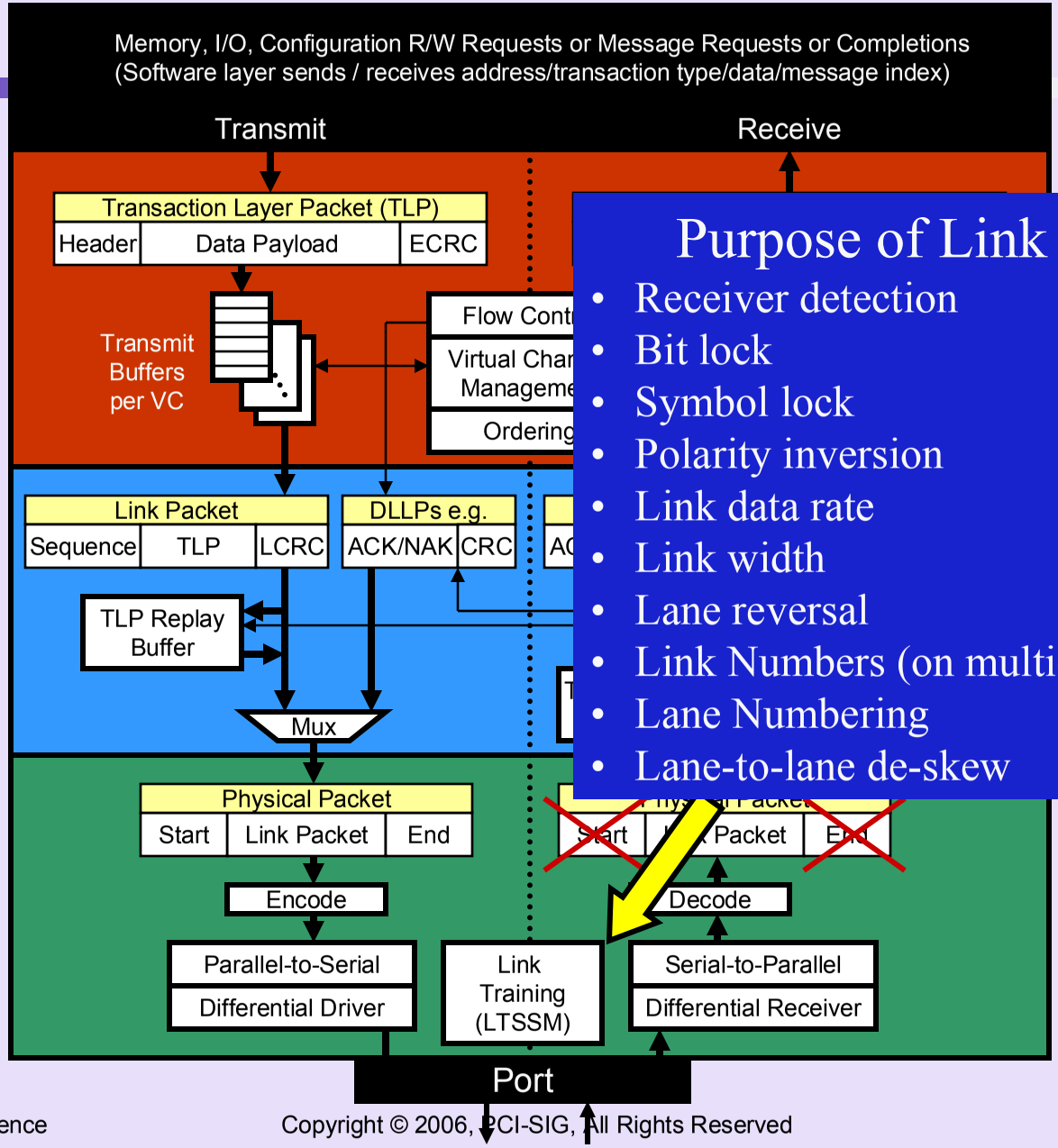
- **Clock driver requirements**
  - ✓ 100MHz with SSC support (e.g. CK410)
  - ✓ Choose low jitter components
  - ✓ System board terminations only



Agenda

- Logical Physical Layer
- Electrical Physical Layer
- Link Training and Initialization (LTSSM)
- Layout Considerations

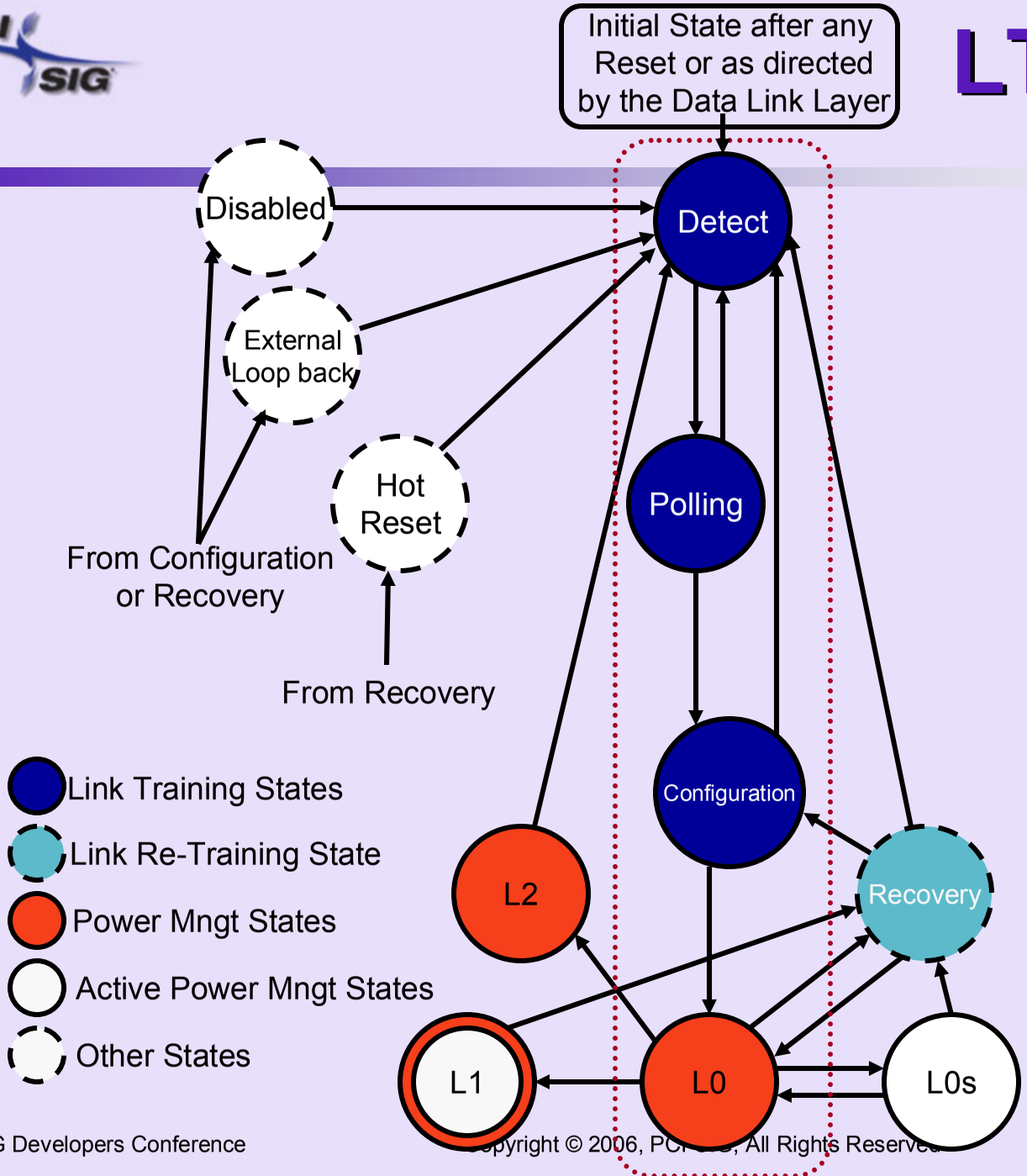
Link Initialization and Training



Purpose of Link Training

- Receiver detection
- Bit lock
- Symbol lock
- Polarity inversion
- Link data rate
- Link width
- Lane reversal
- Link Numbers (on multi-Link device)
- Lane Numbering
- Lane-to-lane de-skew

LTSSM

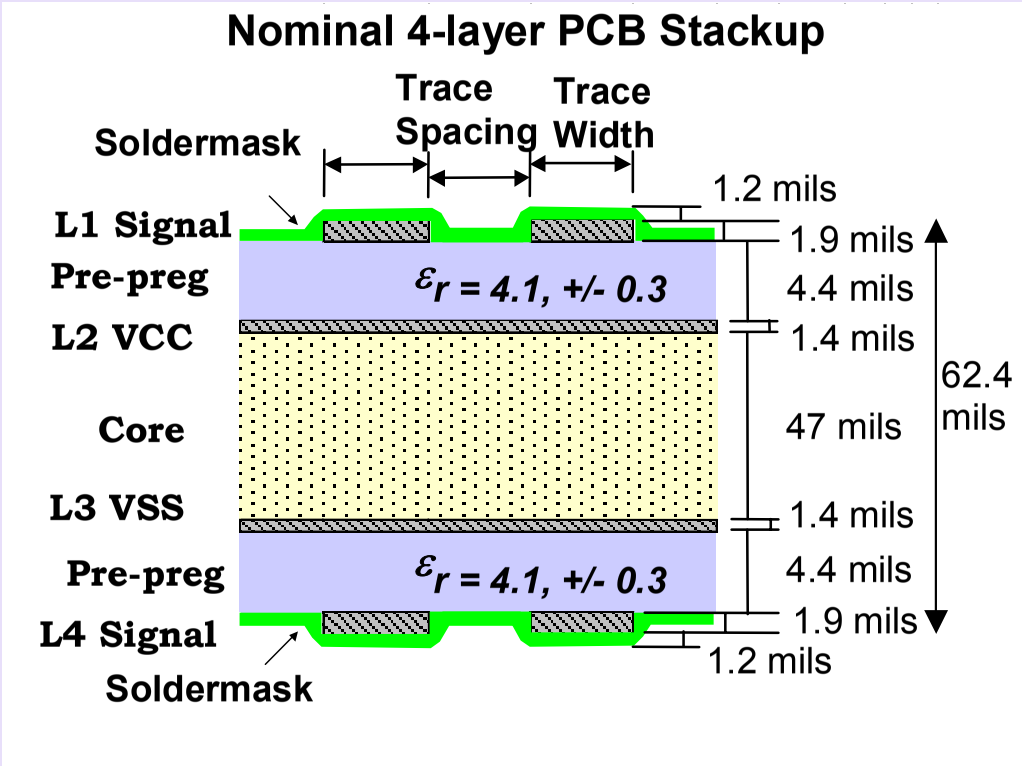


Agenda

- Logical Physical Layer
- Electrical Physical Layer
- Link Training and Initialization (LTSSM)
- Layout Considerations

Stackup design

- No new PCB technology required
- Standard 4-layer stackup 0.062” thick PCB
- Microstrip ½ oz Cu plated Ok
- Stripline 1 oz Cu (6+ layers) Better



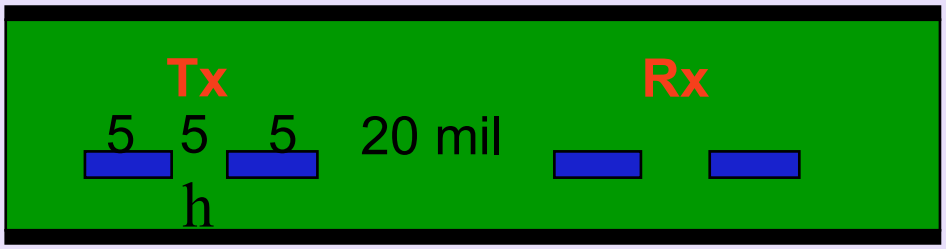
Follow simple layout rules & design tradeoffs

Trace Geometry & Impedance

- Use wider trace width ⇒ Minimize loss
- Use wider traces for long routes
- More pair-to-pair spacing ⇒ Minimize crosstalk
- Target differential Z_o of 100 Ω ±20%

Microstrip

Non-interleaved Topology example

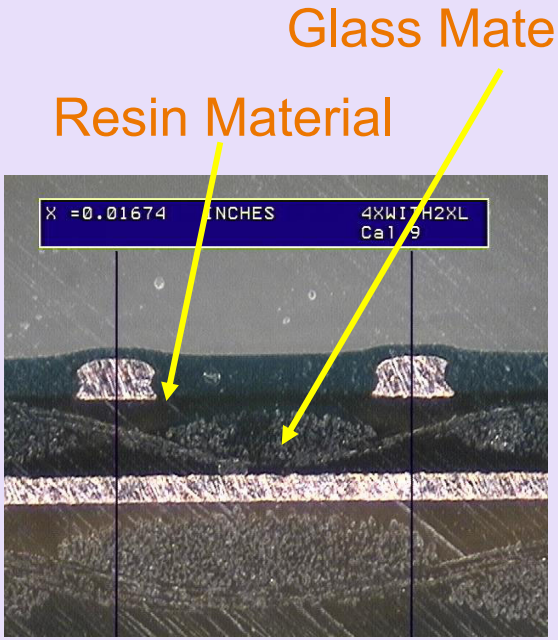


Stripline

Interleaved Topology example

PCB material dominates loss

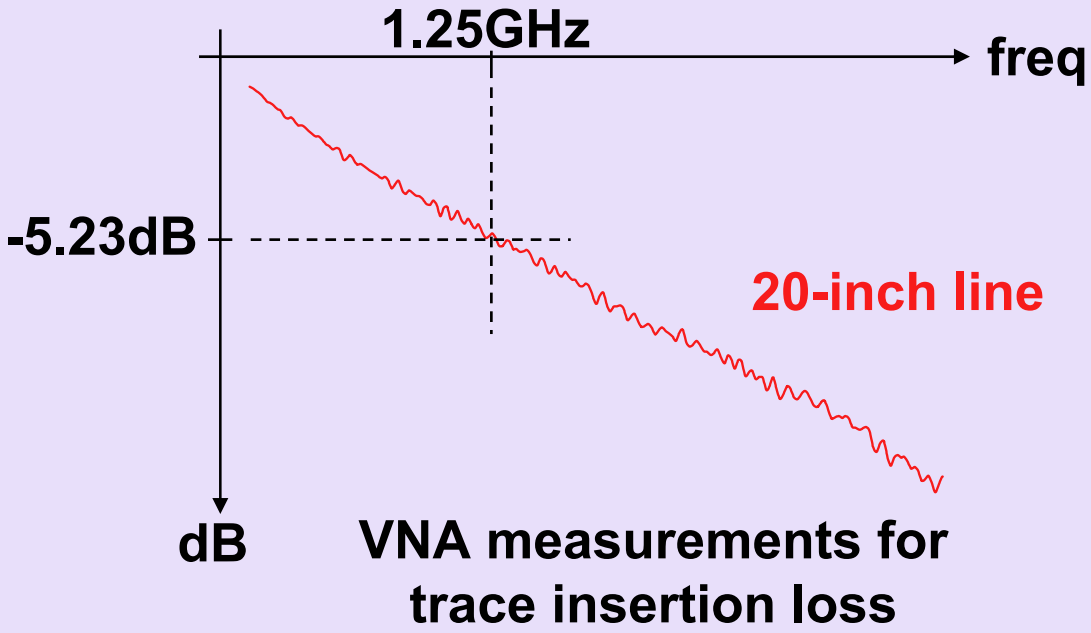
- Stackup FR4 material
  - ✓Copper roughness ⇒ loss ↑
  - ✓Thinner dielectrics ⇒ loss ↑
- Non-homogeneous dielectric
  - ✓Localized Zo variation due to material weave ⇒ loss ↑
- Wide differential Impedance variation on μstrip
  - ✓Etching and Plating process ⇒ loss ↑



FR4 cross-section

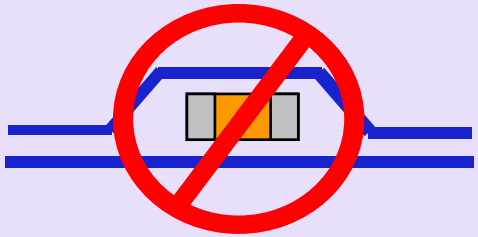
Trace length

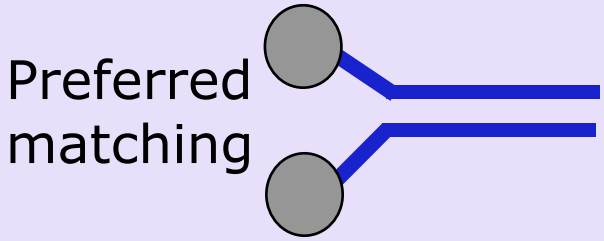
- Longer trace length ⇒ loss ↑
- 0.25 to 0.35 dB inherent loss per inch for FR4 microstrip traces
- Limit motherboard trace to < 12 inches and add-in card trace to < 3 inches

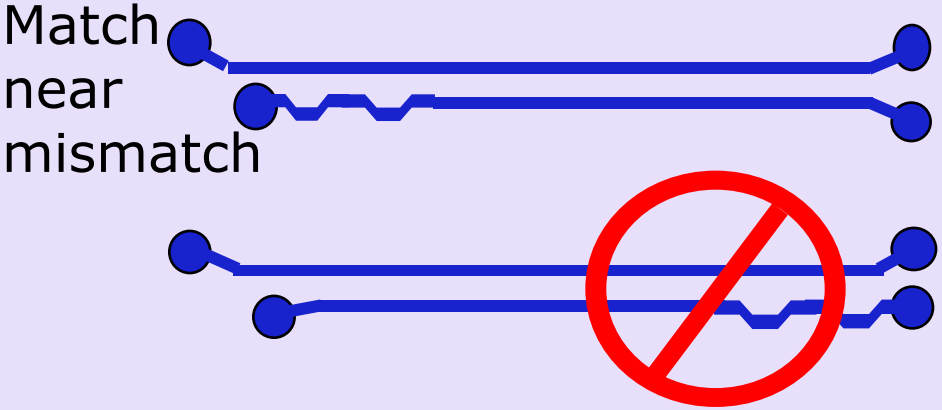


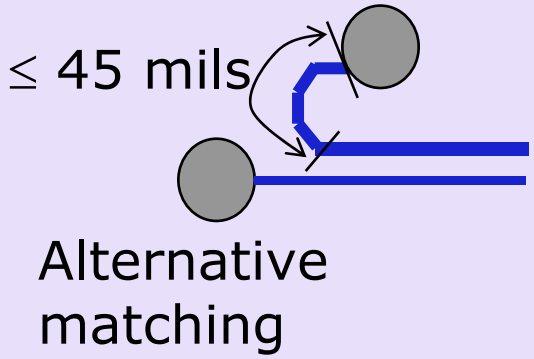
Trace Symmetry & Matching

- Match each differential pair per segment
  - ✓ Match overall length ≤ 5 mils
  - ✓ Symmetric routing for each pair



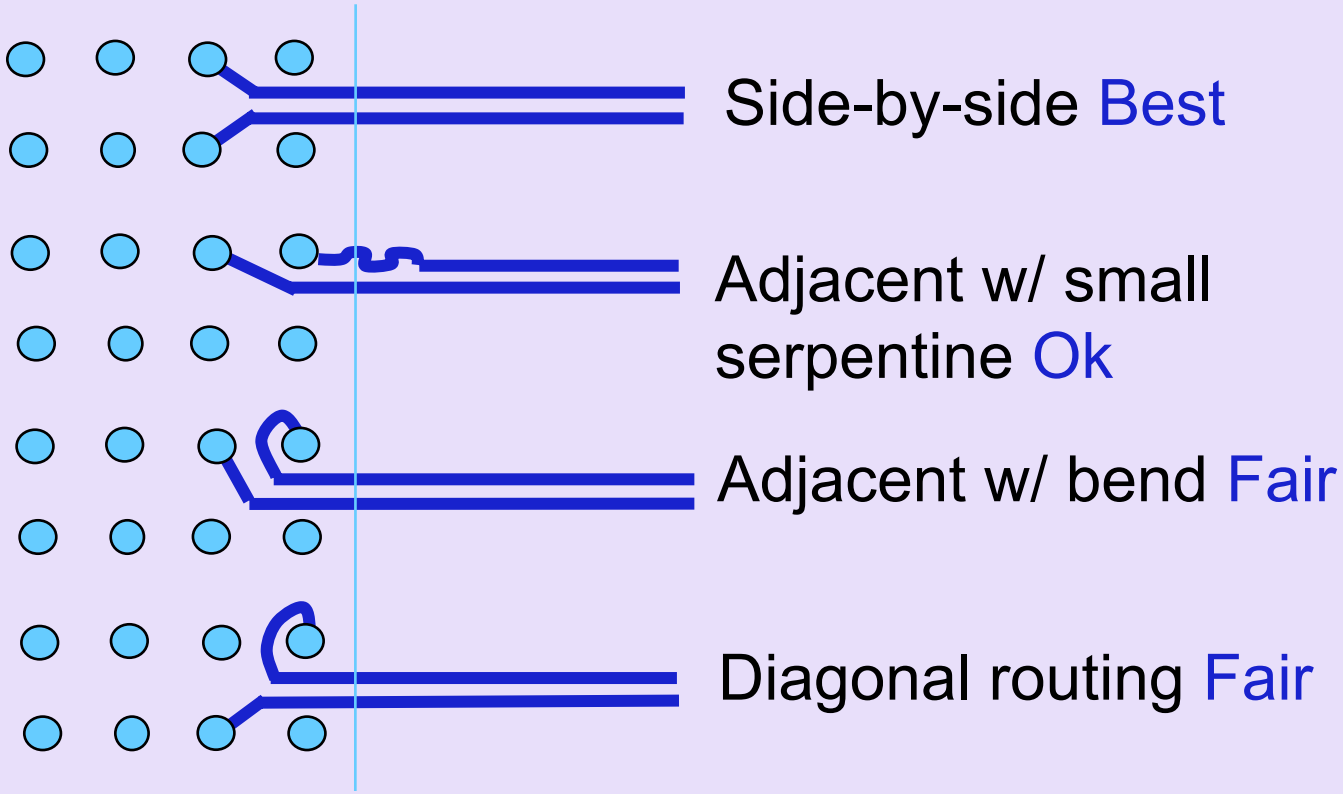






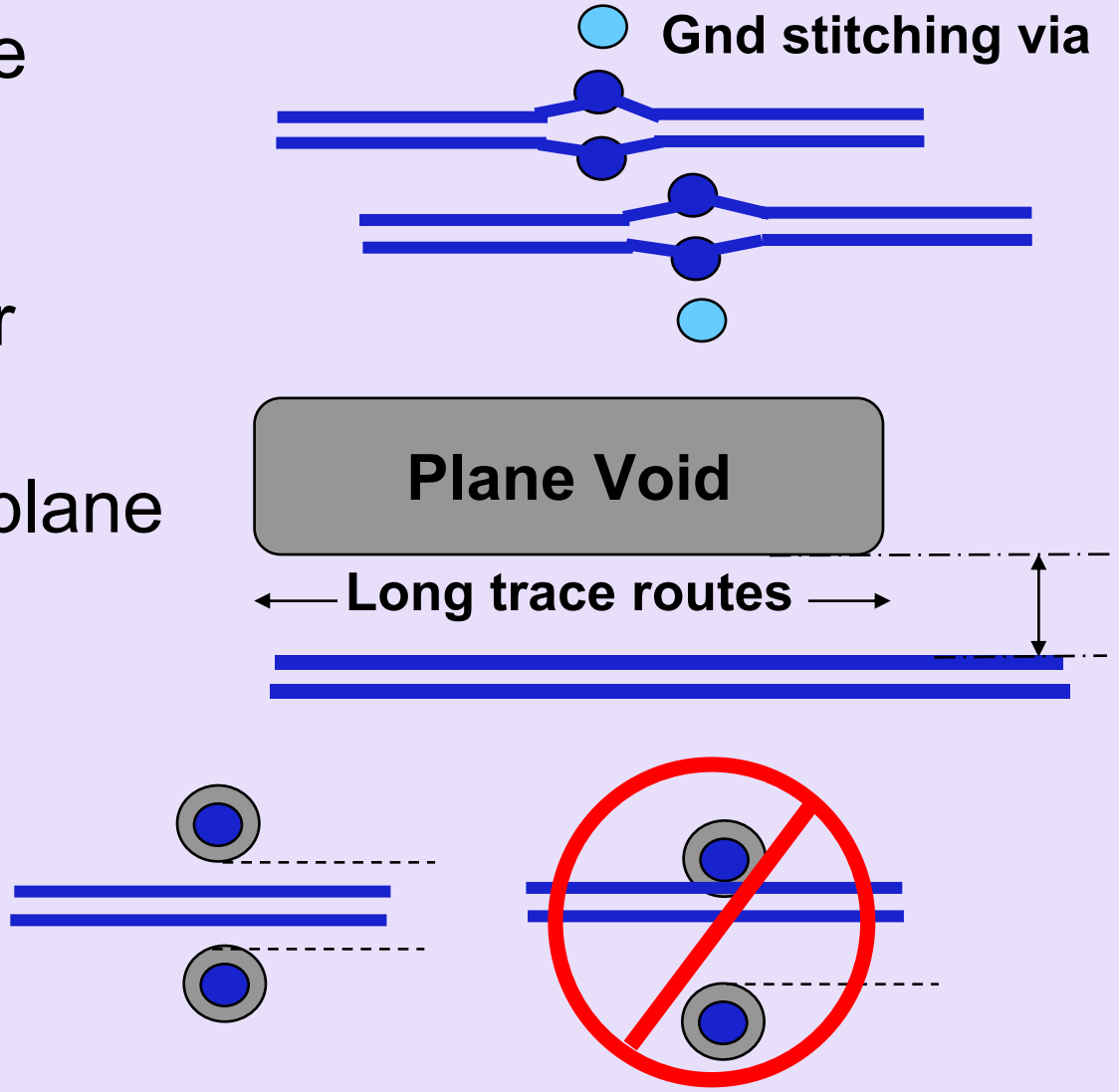
Pin field breakout

- Use side-by-side breakout for package to maintain symmetry
- Avoid tight bends



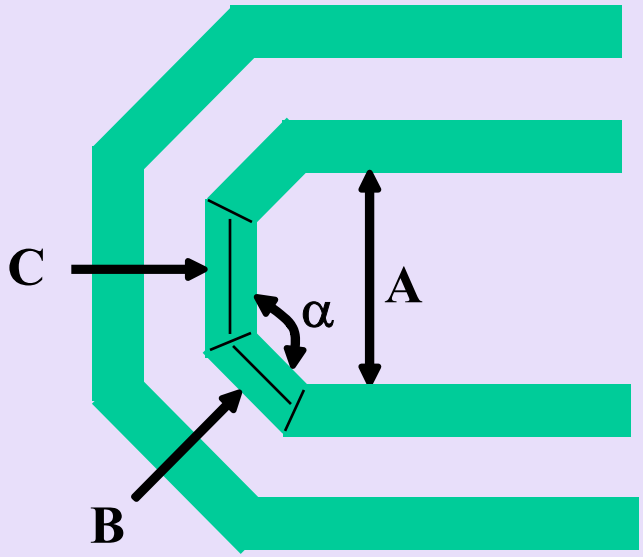
Reference plane

- Full ground plane reference
- Stitching vias required for layer transition
- Clearance near plane void
- Avoid trace over anti-pad



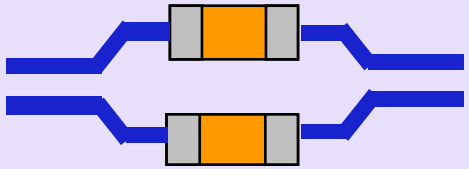
Bend Guidelines

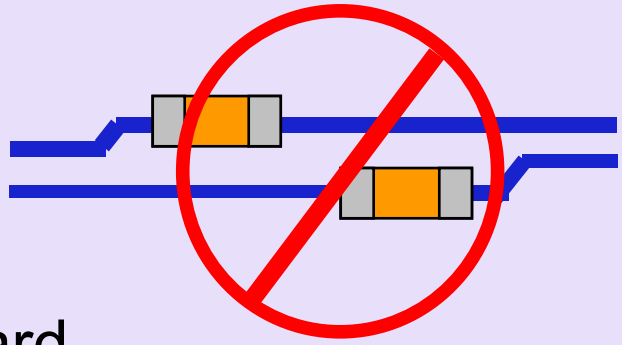
- Avoid tight bends
  - ✓ No 90° bends; impact to loss and jitter budgets
- Keep angles >= 135° (α)
- Keep minimum air gap
  - ✓ A >= 3x the trace width
- Length of B and C >= 1.5x the width of the trace



AC coupling caps

- Size: 0402 best, 0603 ok
- No 0805 size or C-packs
- Symmetric placement

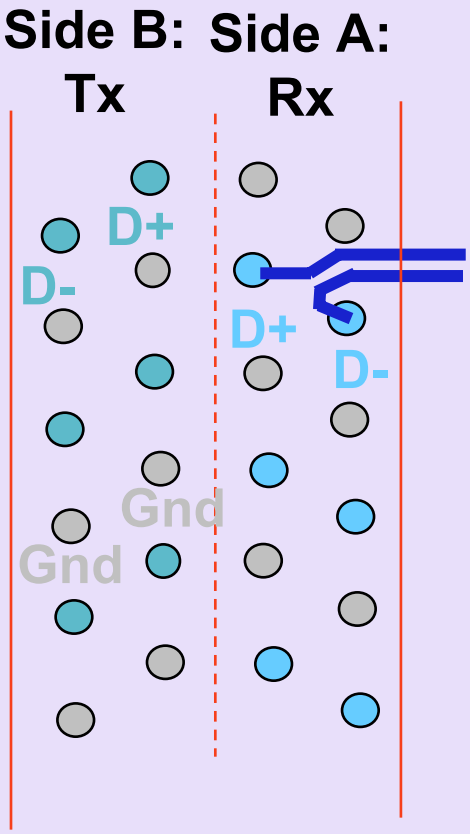




- Cap Size: 0.1uF best
- Cap location:
  - ✓ Along Tx pairs on Motherboard
  - ✓ Along Tx pairs on Add-in card

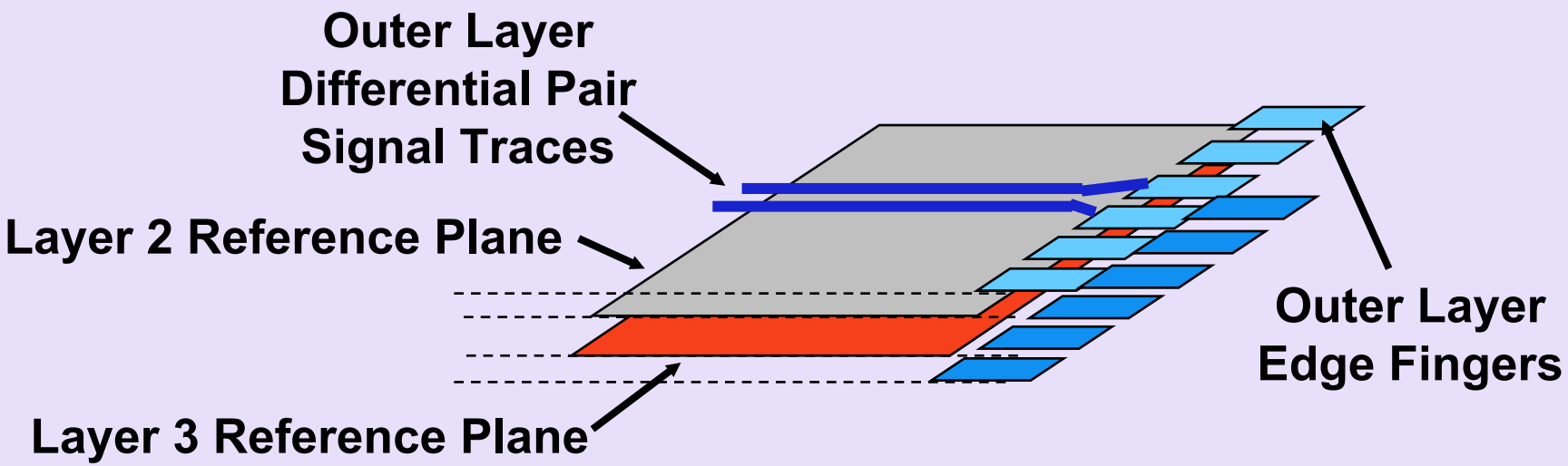
Connectors

- New connector with standard PTH
- Pinout optimized for differential routing
- Loss & crosstalk part of baseboard budget
- Connector sizes: x1, x4, x8, x16



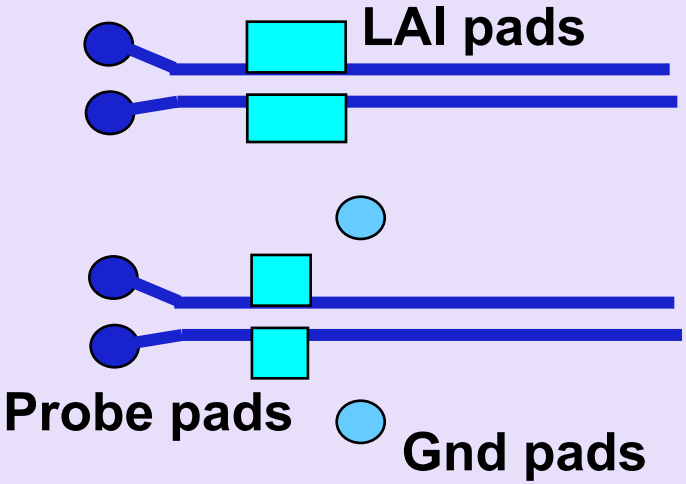
Card edge fingers

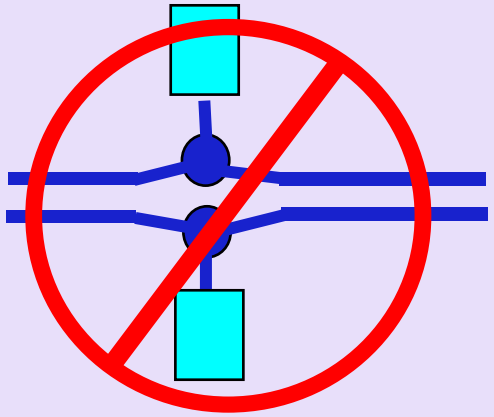
- Remove ref plane under edge fingers pads
  - ✓ For better impedance/loss performance



Test points & Vias

- Minimize Vias usage
  - ✓ Up to 0.25 dB loss per via
  - ✓ Via pad size ≤ 25 mil, hole size ≤ 14 mil
- Put test points or LAI pads in series
  - ✓ No stubs
  - ✓ Provide Gnd pads for single-ended probing





Thank you for attending the PCI-SIG Developers Conference 2006.

For more information please go to

www.pcisig.com

and

www.mindshare.com

PCI Express 1.1 PHY Design Considerations

Milpitas, CA August 21, 2006

Joe Winkles

Staff Architect, MindShare, Inc.