A “Subset PHY” Approach for Energy Efficient Ethernet

Scott Powell, Wael Diab Broadcom

Portland, Oregon Jan, 2008

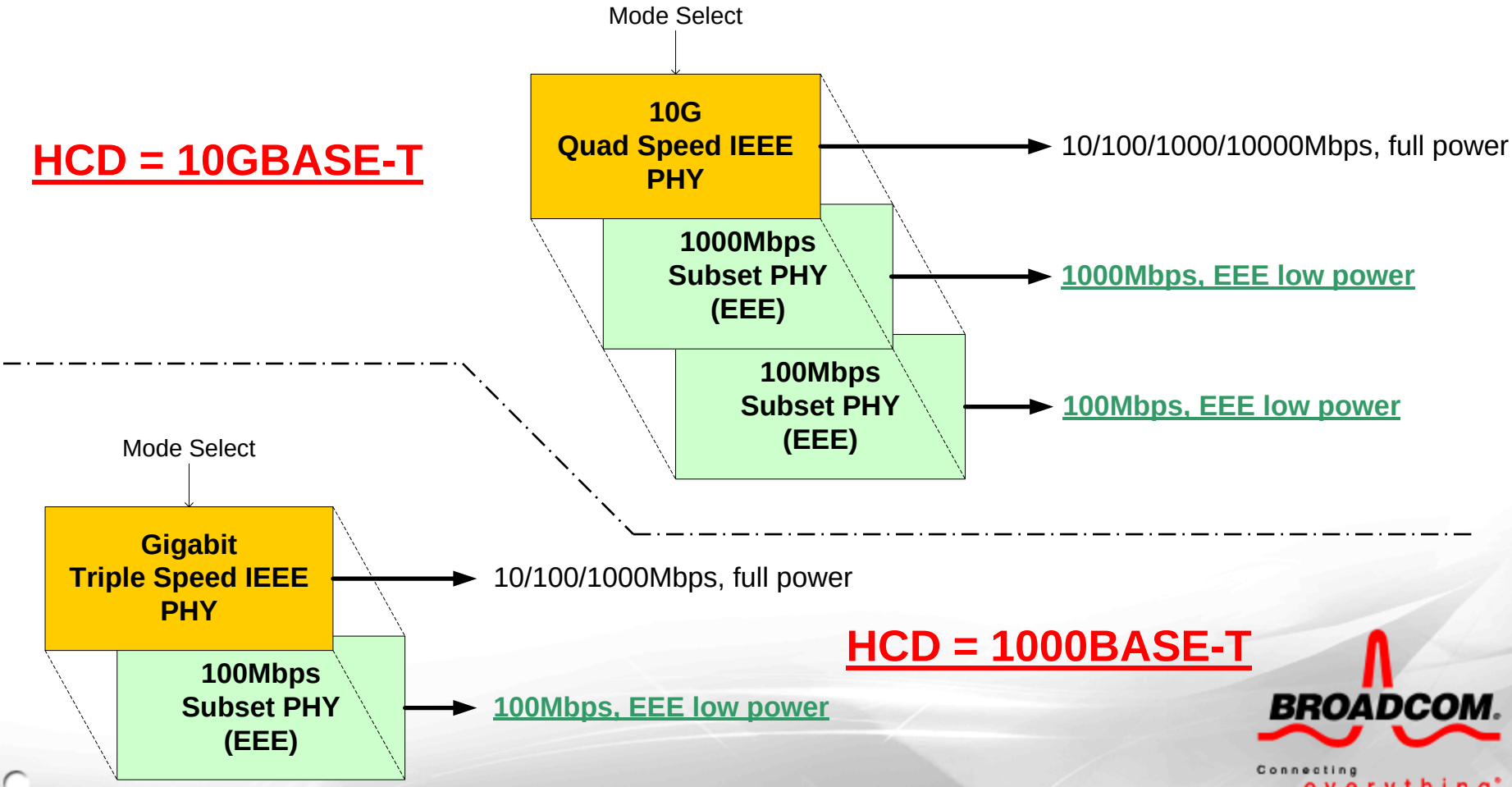
Subset PHY Review

- 802.3 defines a set of distinct copper PHYs for single speed operation
  - Ability to rapidly switch between PHYs not considered
  - Alternative is to define a set of line codes specifically designed to support rapid switching between data rates
- Define a lower rate line code for to be a simple subset of the higher rate line code
  - Highest data rate PHY is a standard *n*BASE-T IEEE PHY
  - Subset PHY implemented by simply turning off elements of HCD PHY
  - A subset PHY is designed specifically to permit rapid speed change
  - Lower data rate subset PHYs designed to retain information necessary for rapid switching back to highest data rate
- Micro-second order of magnitude data rate transition time technically feasible

BROADCOM
Connecting everything

Subset PHY Types

- Define optimal set of subset PHYs for each HCD
  - Results in 3 unique subset PHY types
  - Can go directly from 10G to 100M



Subset PHY Types: Observations

- For the 10GBASE-T HCD, can move between any of the modes
  - For example can go directly from 10G to 100M Subset
- Subsets optimized for Quad Speed and Triple Speed Architectures
- Triple-speed PHYs
  - Largest initial market
  - Triple-speed PHYs on *nth* generation and optimized for technology / process. Thus subset for 100M would optimally be based on 1000BASE-T
    - Overhead to implement subset marginal
- Quad-speed PHYs typically implemented as a Triple Speed Core + 10G Core
  - When HCD = 10G, maintaining different subsets
    - Eliminates interaction with triple speed core
    - Allows “drop in” architecture approach to persist

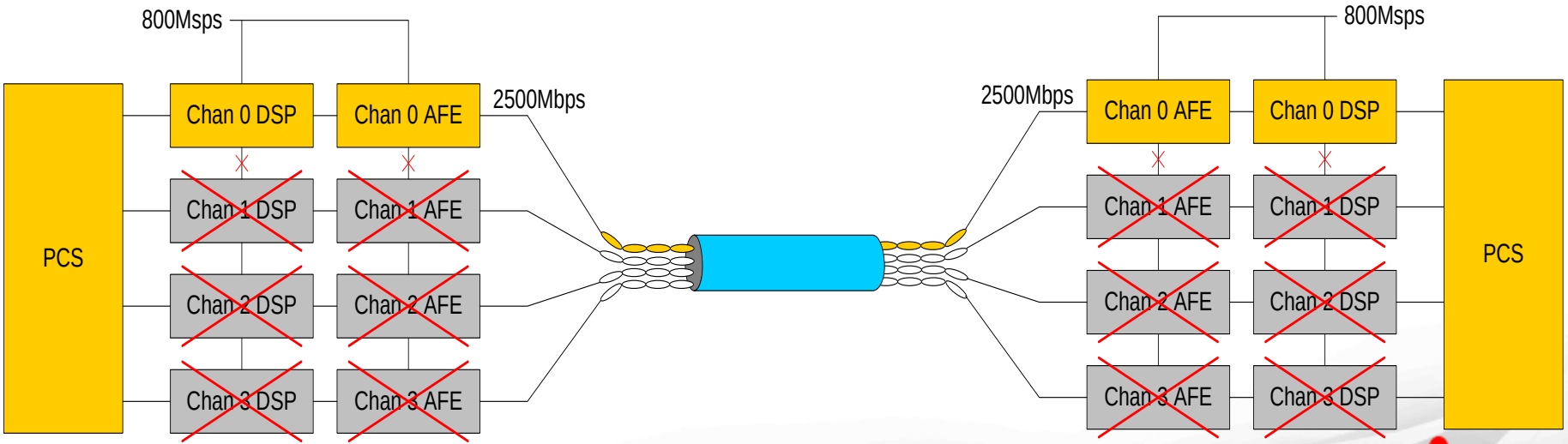
Note: HCD stands for Highest Common Denominator (Clause 28)

BROADCOM
Connecting everything

Subset: Turn off Unused Channels

1a) Turn off 3 of the four channels for 10GBASE-T

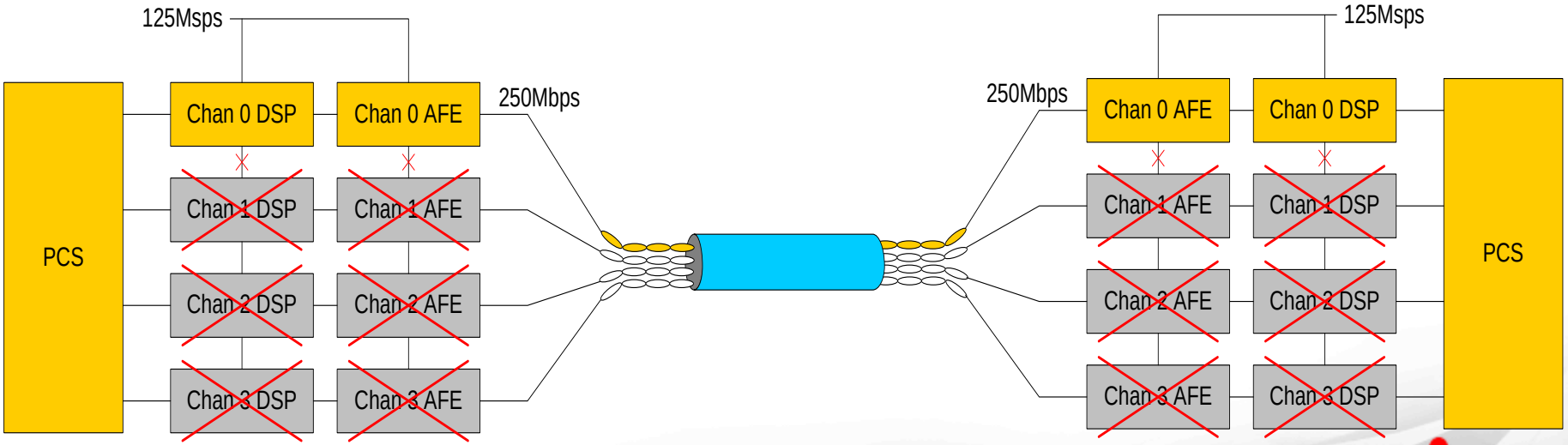
- 800Msps symbol rate remains the same for the active channel
  - Equalizer and echo canceller continue to adapt for active channel
- Synchronization and master/slave loop timing retained
- First step in reducing rate to 1Gbps or 100Mbps



Subset: Turn off Unused Channels

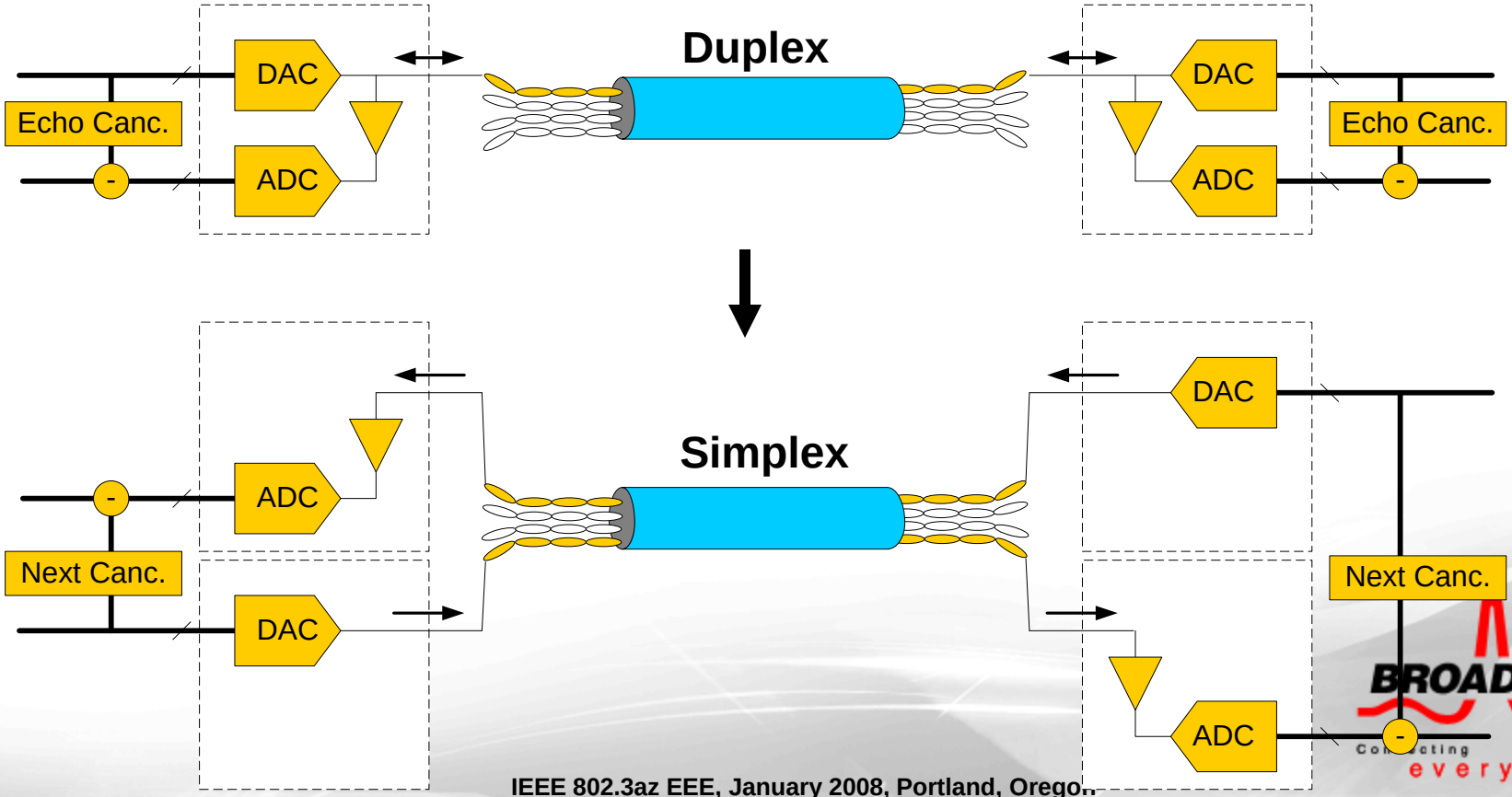
1b) Turn off 3 of the four channels for 1000BASE-T

- 125Msps symbol rate remains the same for the active channel
  - Equalizer and echo canceller continue to adapt for active channel
- Synchronization and master/slave loop timing retained
- First step in reducing rate to 100Mbps



Subset: Duplex to Dual Simplex

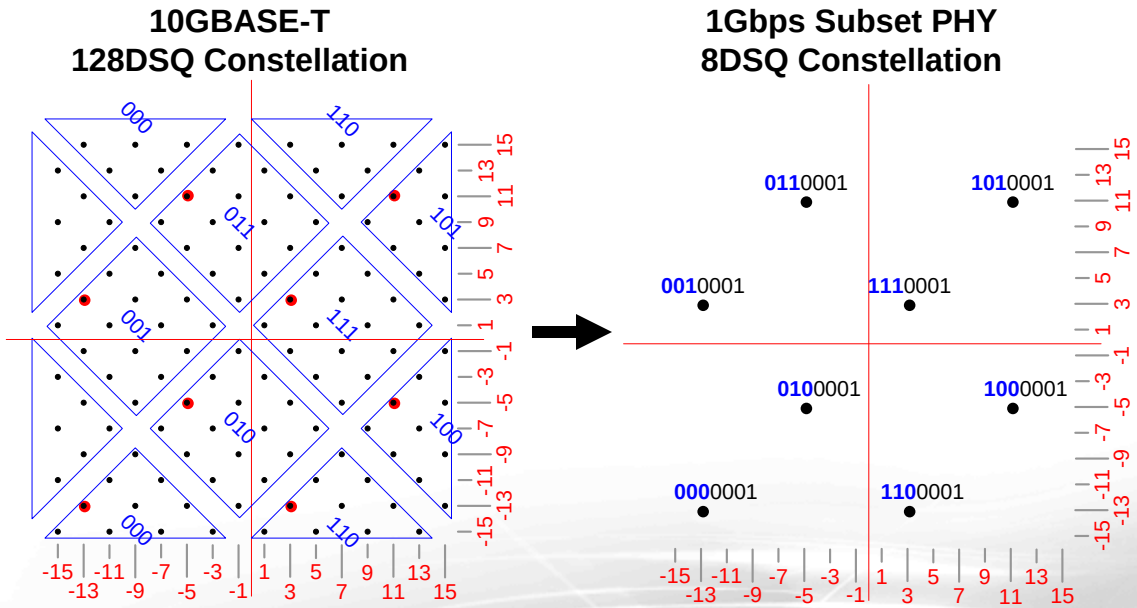
2. Turn off echo cancellers and hybrids for 10GBASE-T or 1000BASE-T
   - Next canceller is usually less complex
   - Better dynamic range on ADC (no echo)
   - Does not change the number of active transmitters or receivers



10GBT to 1G: Reduce Constellation

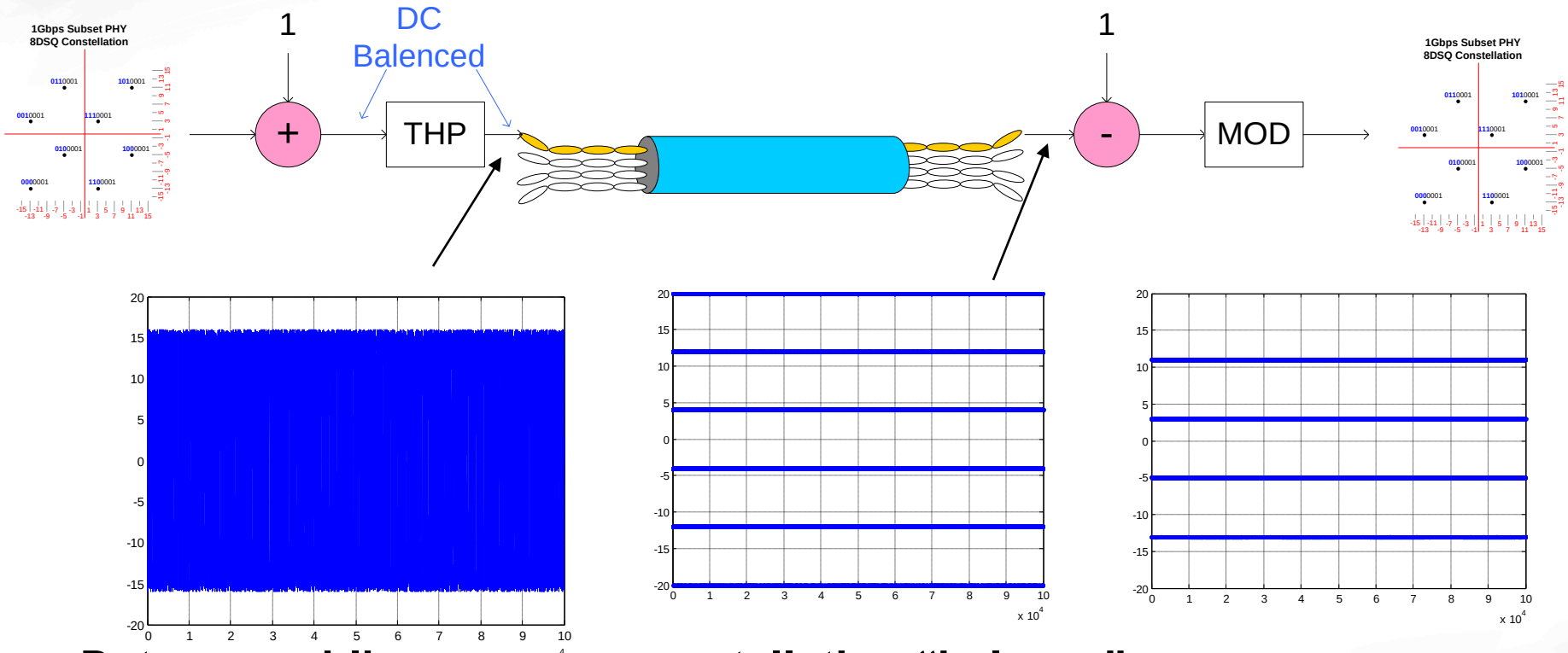
3a) Turn off extra constellation points for 10GBASE-T, 1Gbps mode

- 8DSQ symbols formed by 10GBASE-T 128DSQ subset partitions
- Turn off LDPC encoder and decoder (big power savings)
- Data protected by subset partition gain
  - Same BER as full rate 10GBASE-T
- Enough "spare" bits to add CRC error detection and OOB channel



- 8 Subsets (uncoded)
- 16 points per subset (coded)

Impact of Constellation "Offset"



- Data scrambling ensures constellation "balance"
  - Equal number of positive and negative symbol values
  - Simple DC offset, does not result in "spurs"
  - Can be corrected by adding/subtracting '1'

10GBT to 100M: Reduce Constellation

3b) Turn off additional constellation points for 10GBASE-T, 100Mbps mode

- Reduce constellation to PAM-2 symbols
- Turn off LDPC encoder and decoder
  - Data protected by (at least) subset partition gain
- Turn off pre-coder, use multiplierless DFE
  - Mode already available for PMA training phase
  - Error propagation not an issue (no coding)
  - No constellation expansion
- Reduced transmit power (simple pam-2, not uniform)
  - 3x reduction in transmit amplitude, ~9x reduction in transmit power
- Eliminate all multipliers from cancellation filters
  - Major source of power dissipation
- Enough "spare" bits to add CRC and OOB channel for 10G upshift control

BROADCOM
Connecting everything

1000BT to 100M: Reduce Constellation

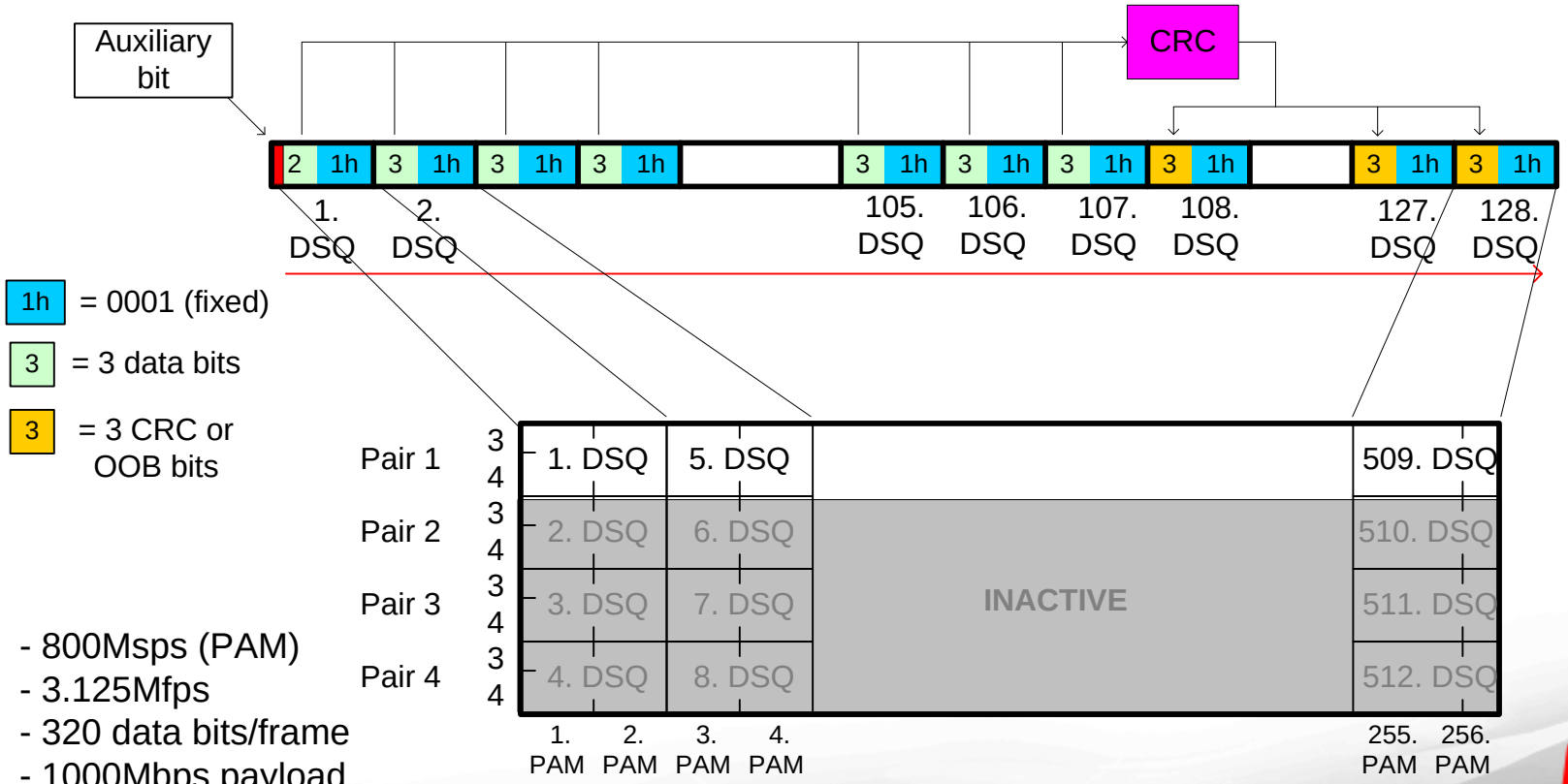
3c) Turn off extra constellation points for 1000BASE-T

- Reduce constellation to PAM-2 symbols
- Turn off convolutional encoder and Viterbi decoder
  - Data protected by (at least) subset partition gain
- Eliminate multipliers in DFE
  - Error propagation not an issue (no coding)
- Reduce transmit power
  - 1.5x reduction in transmit amplitude, ~2x reduction in transmit power
- Turn off partial response filter
  - Not necessary for EMI due to reduced transmit power
- Eliminate all multipliers from cancellation filters
- Pad symbol rate to 125Msps
  - Define OOB channel
  - Data already scrambled – dc balance and transition density not an issue
  - Can be done multiple ways, for example a simple 8B10B transcode

Potential to reduce power below 100TX

HCD=10G: Subset PHY Framing

- **Simple modification to 10GBASE-T Clause 55 framing:**
  - **Example: 1000Mbps (similar for 100Mbps)**



- 800Msps (PAM)
- 3.125Mfps
- 320 data bits/frame
- 1000Mbps payload

Updating Inactive Channels

- Multiple presentations indicate only infrequent updates needed to maintain state of inactive channels
- OOB channel of subset PHYs can be used to synchronize updates
- Inactive channels and cancellers not being used for anything else
  - Available to be used for refresh operation
  - Does not impact data transfer, link maintained
  - Does not require switching back to higher speed
- Inactive channels can be updated by turning them on for short bursts with long duty cycle
  - Transmit idles
  - Very low duty cycle makes update power negligible

BROADCOM
Connecting everything

Synchronizing Transition Between HCD and Subset PHY Modes

- Synchronization done at the PHY layer
  - Simplify higher layers
  - Exploit existing master-slave timing relationship
- Downshift: Transitioning from HCD (standard IEEE) mode to subset PHY EEE mode
  - Synchronization at packet boundaries is sufficient
    - Not much packet transfer preceding a downshift by definition
    - Relatively insensitive to speed of transition
  - Downshift can be synchronized by newly defined Q-ordered set in the IPG idle sequences
    - Both link partners must be EEE capable -> both link partners will have newly defined Q-ordered set
    - Either unused available Q-ordered sets or use of “illegal” idle symbols
- Upshift: Transitioning from subset PHY EEE mode to HCD mode
  - Very sensitive to speed of transition
  - Upshift can be rapidly synchronized by OOB channel

BROADCOM. Connecting everything®

HCD=10G: Subset PHY Power Estimate

- Analog (1000Mbps)
  - 3 of 4 ADCs and DACs powered down
  - All 4 hybrids powered down
  - Estimate 85% power saving (1000Mbps)
- Digital (1000Mbps)
  - All 4 echo cancellers powered down
  - 11 of 12 NEXT cancellers powered down
  - All 12 FEXT cancellers powered down
  - 3 of 4 equalization filters powered down
  - 3 of 4 Precoder filters powered down
  - LDPC encoder and decoder powered down
  - Estimate >85% power savings (1000Mbps)

- Analog (100Mbps)
  - Above + reduce transmit power
  - Estimate >33% additional power saving (100Mbps)
- Digital (100Mbps)
  - Above + multiplierless filters
  - Estimate >40% additional power savings (100Mbps)

- Estimate: ~85% power reduction in 1G subset PHY mode
  ~35% additional reduction in 100M subset PHY mode

HCD=1G: Subset PHY Power Estimate

- Analog
  - 3 of 4 ADCs and DACs powered down
  - All 4 hybrids powered down
  - Reduce transmit power
  - Estimate 50% power saving
- Digital
  - All 4 echo cancellers powered down
  - 11 of 12 NEXT cancellers powered down
  - 3 of 4 equalization filters powered down
  - 3 of 4 DFE powered down
  - Viterbi decoder powered down
  - Estimate >75% power savings

- Estimate ~65% power reduction in 100M subset PHY mode
  - Potentially less than 100BASE-TX

BROADCOM
Connecting everything

Summary

- Optimized Subsets proposed for 10GBASE-T & 1000BASE-T PHYs
  - Minimal control enhancements to base PHY required to implement Subsets
  - Simple approach of “turning off” blocks in the HCD PHY proposed
- Coefficients always fresh
  - Eliminates possibility of stale coefficients / non-convergence
  - Active elements always fresh
  - Simple low-cost mechanism proposed to keep coefficients fresh on silent elements
- Very rapid speed shifts attainable with high power savings
  - Micro-second order of magnitude data rate transition time technically feasible
  - ~7x or greater power reduction for 10G to 1G
  - Lower power for 100M than 100BASE-TX
- Precise communication and control possible between link partners
  - Take advantage of master-slave relationship. Link partners perfectly synchronized
- Very rapid Upshift feasible through use of OOB
  - Buffering determined by upshift time
  - Allows for low buffer systems to be implemented

BROADCOM
Connecting everything