A “Subset PHY” Approach for 10GBASE-KR Energy Efficient Ethernet

Vivek Telang, Velu Pillai, Broadcom

IEEE 802.3az, Orlando, Florida, Mar 2008

Outline

- Section I
  - Answers to questions asked in January meeting
  - More details on the BP Subset PHY Proposal
- Section II
  - Review of January presentation

Equalizer Refresh

Q: Are the Eq coefficients always fresh when signaling at 1/10 rate?

- The low frequency equalizer coefficients will stay fresh at the 1G rate
- But the higher frequency dimensions of the equalizer will not be excited by a 1G signal, and will need a periodic refresh at the 10G rate.
- The refresh period can be of the order of 10s of seconds, maybe minutes – further testing is required
  - This will have minimal impact on the energy efficiency
- Note that since the rate switching is synchronous, the parent clock is always tracked, and clock frequency drift is not an issue

Lower Rate Equalization

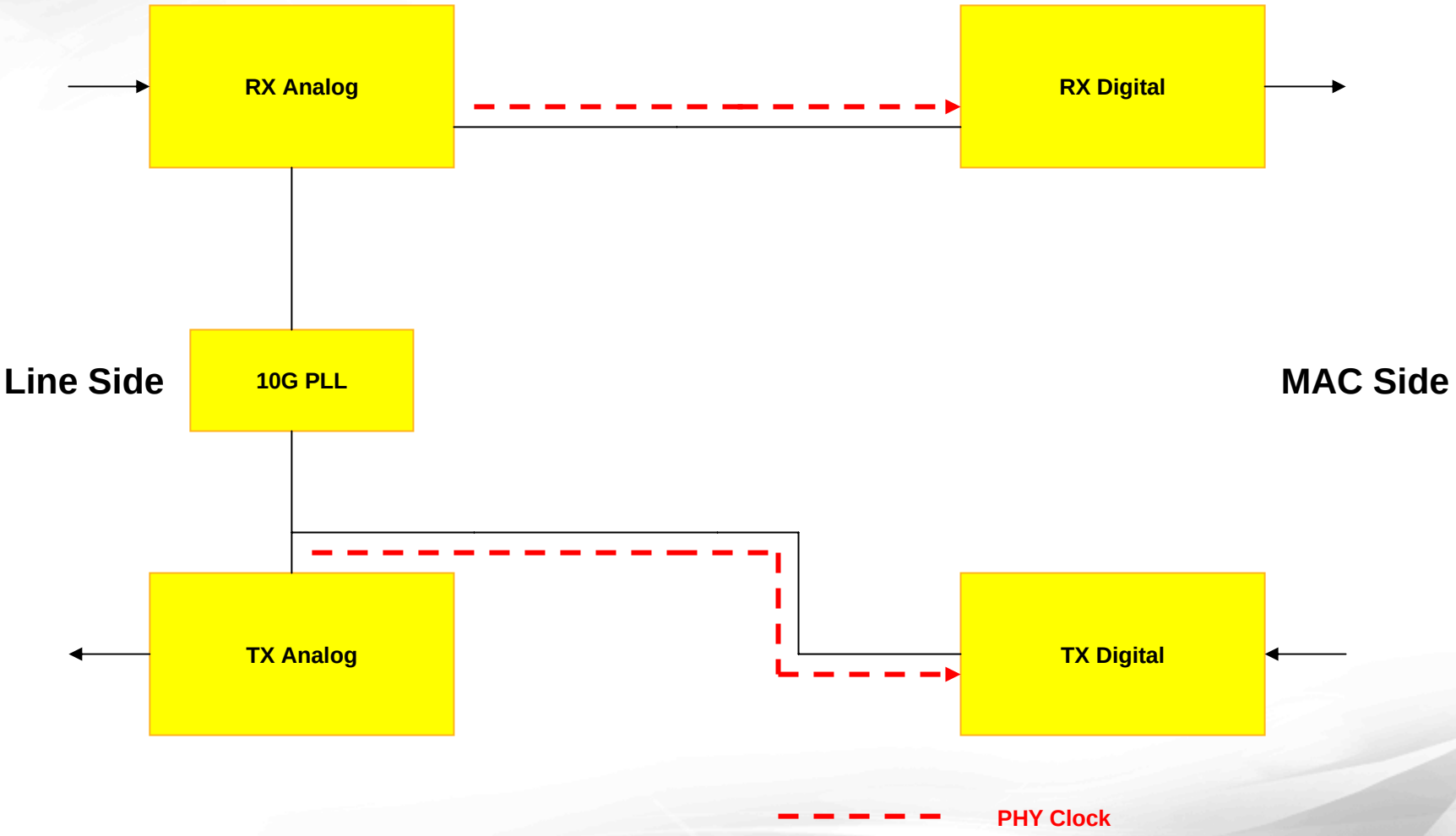
Q: Can the lower rate equalizer be derived from the higher rate Eq?

- Preliminary analysis indicates that 1G signal does not require equalization
  - For example, Rx eye is open at the output of 1000BASE-KX channels
- If required, lower rate Eq setting can always be derived from the higher rate equalizer. For example,
  - Analog Rx filter architecture
    - Filter H(s) is the approximate inverse of channel transfer function, independent of rate
    - No change required at lower rate
  - Analog DFE architecture
    - Analog Filter H(s) stays the same
    - 1G DFE coefficients can be copied from 10G DFE and decimated by 10
  - Digital FFE/DFE architecture
    - FFE and DFE coefficients can be copied from 10G, and decimated by 10

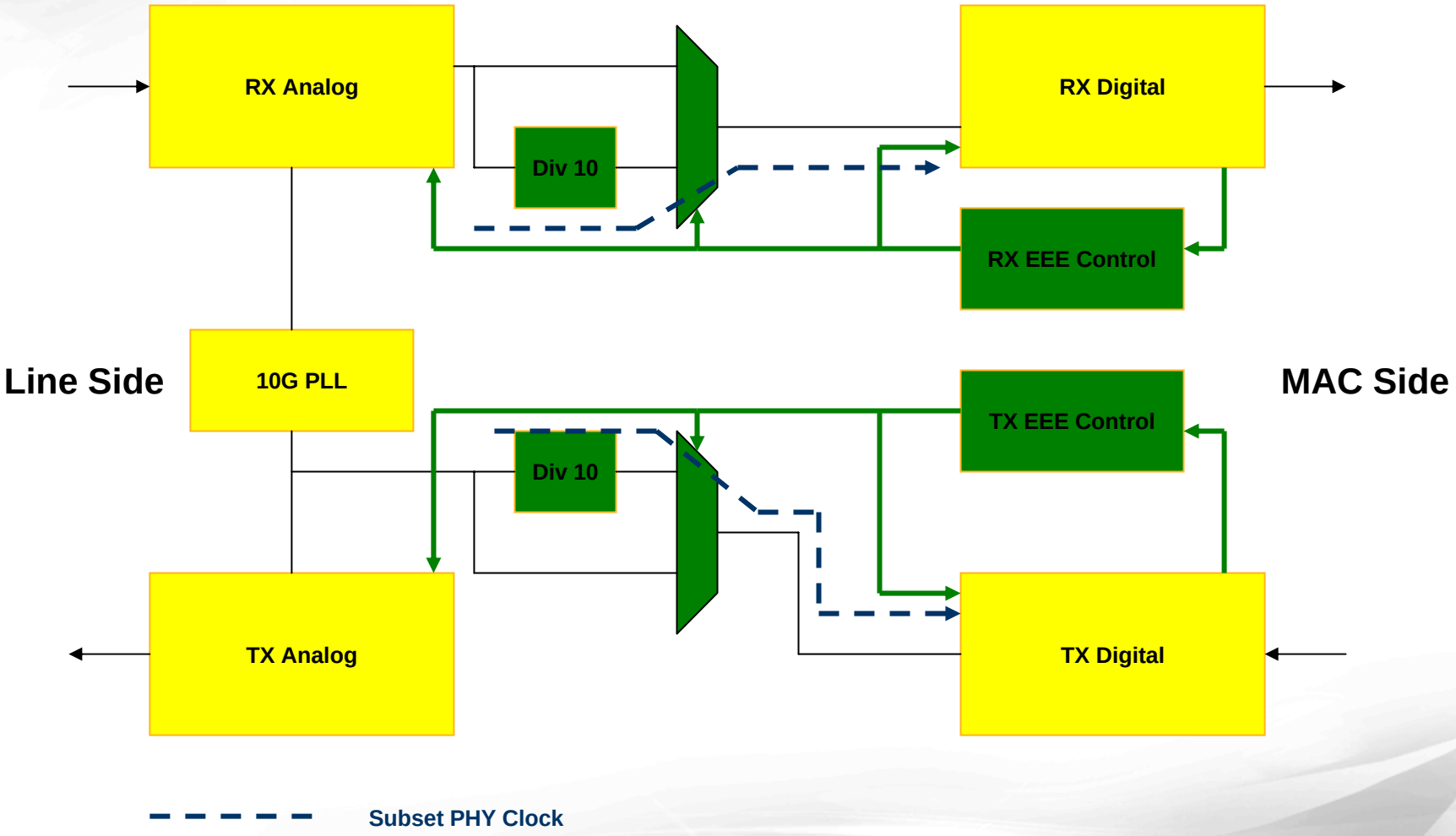
Synchronous BP Subset PHY Clocking

- PLL clock stays the same
  - 10G clock and 1G clock are source synchronous.
- Analog blocks always operate on the 10G clock, no switching issues
  - Symbol rate is reduced by 10.
  - 10G blocks may be powered down in 1G for power savings
  - During upshift, Programmable Byte count ensures analog power up settling time
    - Max and min values to be determined by analog simulations
- Digital blocks are clocked at either full rate or 1/10th rate
  - Synchronous clock switching circuitry guarantees smooth transition during up or down shift between 10G and 1G rates
  - Exchange of transition symbol ensure the precise switching clock cycle for all the down stream digital logic.

Sample Implementation of 10G KR PHY



Sample Implementation of Synchronous Subset PHY Rate Transition

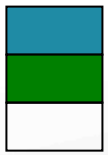


Power Savings Assumptions

- Power analysis based on example architecture described below
  - Idea was to abstract out implementation details and provide a 1st order analysis
  - Exact power savings may change with specific architecture and process geometries
- Example Architecture Assumptions
  - Analog
    - TX Pre-Emphasis and Line Driver
    - VGA, RX equalizer, PLL
  - Digital
    - Shared logic (MDIO registers, diagnostics, etc.)
    - TX Coding
    - RX Equalization & Timing Recovery
    - PCS & Auto Negotiation
- When going from 10G to 1G, subsystems are either switched off or in a lower power mode
  - Lower power mode may run at 1/10th the clock or involve a subset of 10G block
- Approximate energy savings of 75%

Power Savings, by Block Usage

Block	10G	1G
TX Pre-Emphasis		
Line Driver		
VGA		
Rx Equalizer		
PLL		
Timing Recovery		
CDR		
PCS		
Shared logic		



ON

REDUCED POWER

OFF

Observations

- Only requires simple mux to select between parent clock and divide-by-10 clock
  - Eliminates complexity to go to another parent clock speed
- Allows quick, *synchronous and very precise* transitions
  - Current investigation indicates that the Synchronous clock nature of this technique is technically feasible to allow transitions on byte boundaries
  - Maybe useful in upshifting during max length packets
  - Currently running simulations to verify no impact to BER
  - Welcome input from others on feasibility analysis
- Eliminates complexity of switching between 10G and 1G machines
- Simple and Limited Control Overhead
  - No changes to the wire signaling
  - Marginal addition to the control logic of a 10GBASE-KR device
    - Q-ordered sets already defined for 10GBASE-KR. Logic already there
  - Implementation cost estimate is ~1% of KR PHY
- Transition Times and Power Savings
  - A 4X power savings can be achieved within ~10s of usecs transition times

A “Subset PHY” Approach for 10GBASE-KR Energy Efficient Ethernet

Howard Baumer, Wael Diab Broadcom

IEEE 802.3az, Portland, Oregon, Jan 2008

Overview

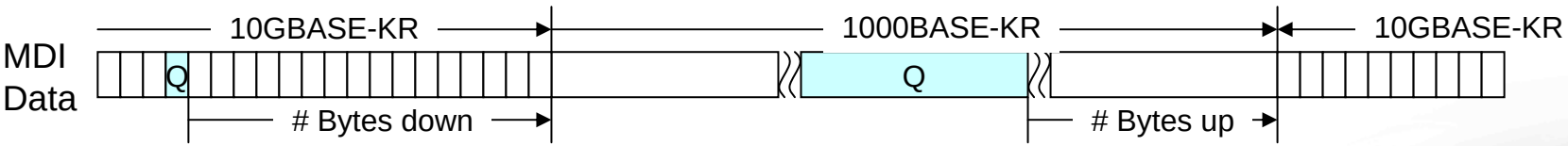
- Objective
  - Identify an approach that will offer fast transition times from the 10G rate to a 1G rate to reduce energy consumption during periods of low link utilization on BP Ethernet KR links
- Two Possible Approaches
  - Rapidly switch between 10GBASE-KR and 1000BASE-KX
  - Define a “Subset” of 10GBASE-KR
  - Subset approach introduced in this presentation
- Line code for lower data rate is a simple subset of the higher data rate (standard) mode
  - Subset PHY implemented by simply turning off elements of higher data rate standard parent PHY
- Communication Method Using Q-Ordered Sets Introduced

10GBASE-KR Subset Phy

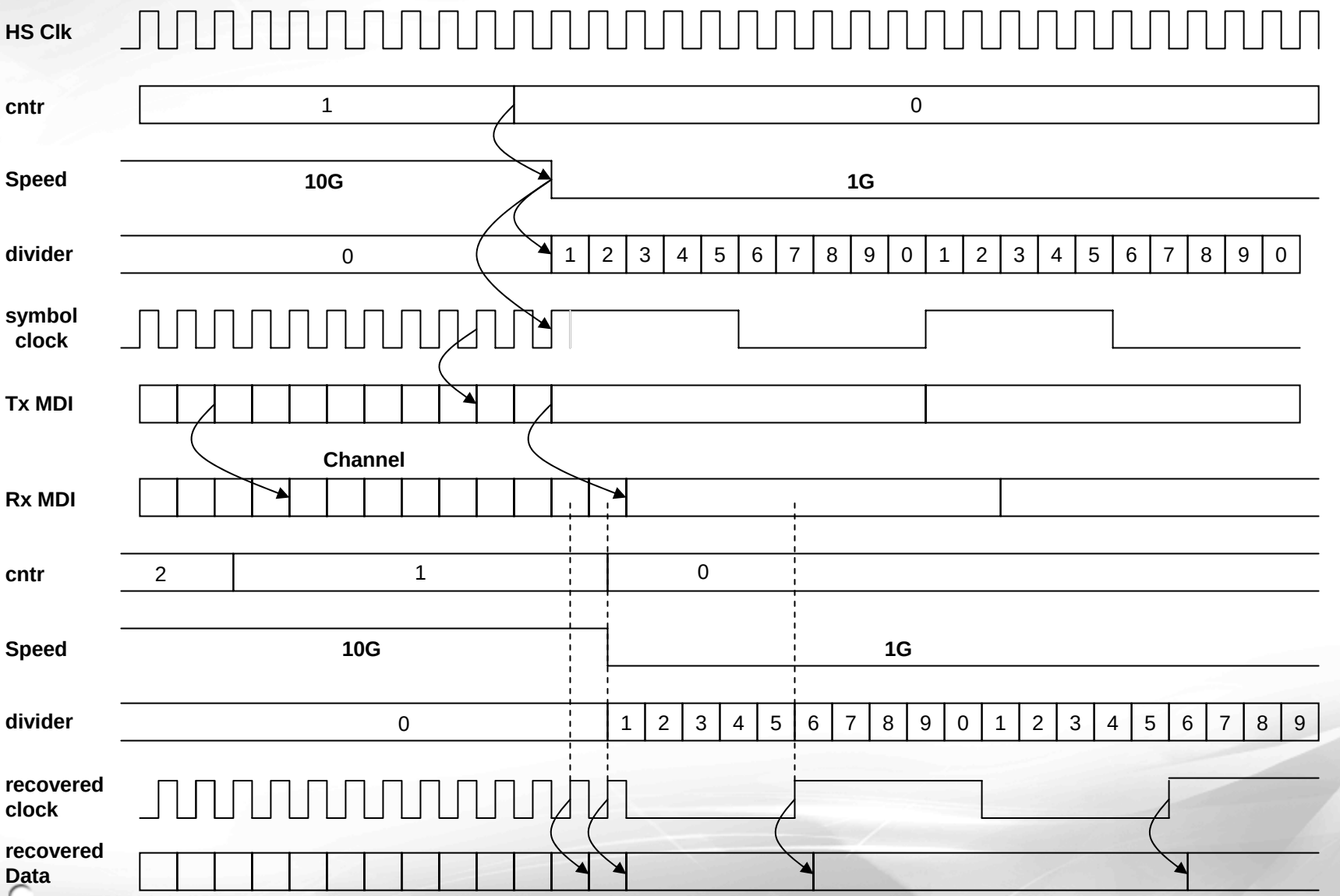
- Define 1 Gigabit Rate as a Subset of 10GBASE-KR by transmitting at 1/10th rate
  - Internal clocks, transmitted symbols, etc. are 1/10th 10GBASE-KR
- Use same 64/66 endec that runs @ 1/10th the rate
- Turn down analog biases, shut off parallel circuits, etc.
- Synchronously change speed up or down
- Allows 64/66 endec to get continuous clocks so it will never loose sync!
- Auto-Negotiate SSP parameters

Precise Transition Communication

- Define ||Q|| to indicate the precise symbol where the speed change will occur
  - $||Q_{dn}||$ = /Q/D0.0/D=down/n/, n=number >2 to be picked
  - $||Q_{up}||$ = /Q/D=up/D0.0/n/, n=number >2 to be picked
- ||Q|| exchanged during IPG



Example Down Shift



Example Up Shift