Efficiency and EEE

Technical Feasibility

Bill Woodruff, Paul Langner and Hossein Sederat

May 29, 2007

Supporters

Scott Powell - Broadcom

Technical Feasibility Using Currently Defined PHYs

A good goal would be to meet the scope using currently defined PHYs:

"...transition to and from a lower power consumption state in response to changes in network demand. The transition will not cause loss of link as observed by higher layer protocols."

- A 10GBASE-T PHY would
  - RPS (Rapid PHY Switching) to 1000BASE-T, then RPS back up to 10GBASE-T
    - Or RPS further down to 100BASE-T, then 10BASE-T
- A 1000BASE-T PHY would
  - RPS to 100BASE-T, then RPS back up to 1000BASE-T
    - Or RPS further down to 10BASE-T
- (if we chose) A 100BASE-T PHY would
  - RPS to 10BASE-T, then RPS back up to 100BASE-T

RPS Issues - Downshifting

Issues in the downshift

- 10GBASE-T to 1000BASE-T
  - Nontrivial. Calculate the equivalent response of many filters (FEXT, NEXT, equalizers, echo) convolved with the 4 THP filters, and then decimate to create starting point for 125MHz 1000BASE-T.
  - CDR must recover clock / phase
  - Adds a math coprocessor, time required is tbd
- 1000BASE-T to 100BASE-T
  - Should be straight forward, same time base
- 100BASE-T to 10BASE-T
  - Similar configurations, revert to Manchester code

RPS Issues - Upshifting

Issues in the upshift

- 10BASE-T to 100BASE-T
  - Should be straight forward, under 1ms
- 100BASE-T to 1000BASE-T
  - Cable temperature change can drive extensive re-training (chadha_1_0407.pdf)
  - 100ms may be seen, implementation specific
- 1000BASE-T to 10GBASE-T
  - Extensive retraining will be required
  - May exceed 100ms, implementation specific
    - One data point of 20ms in zimmerman_01_0307.pdf, we believe is not worst case

Implications w/ Existing PHYs

A worst-case RPS transition time may exceed 100ms

- Transitions “without loss of link” become problematic
- Implications within a multiple segment link amplify effects of this long transition time.

May increase PHY cost

- While most PHYs are multi-rate, none include RPS today
- Adding RPS will add circuit complexity if transition time is to be minimized

Alternative Approach

This approach is intended to address Technical Feasibility

An open discussion is anticipated should a task force be formed, where any number of approaches will be discussed

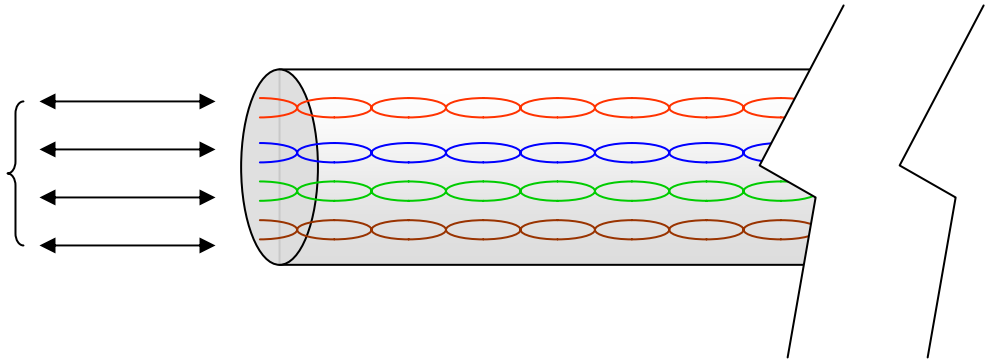
Define a Subset PHY

1. **Set the primary objective as minimizing transition time**
2. **Constrain for minimal systems impact, cost and power**

Can a Subset PHY mode be defined in a 10GBASE-T (or 1000BASE-T) PHY to downshift and upshift in minimal time utilizing elements which are already present?

Current 10GBASE-T Scheme

Bidirectional PAM-16 x4
@ 800 Ms/s with 128-DSQ
(encoding 3.5 bits/symbol)
= 11.2 Gb/s raw bit rate per direction



Elements:

- analog hybrid
- 4x THP
- 4x echo cancellers
- 12x NEXT cancellers
- 12x FEXT cancellers
- 4x Equalizers
- high performance A/Ds and D/As

Possible Subset PHY

Adaptations of 10GBASE-T to operate at GE rates

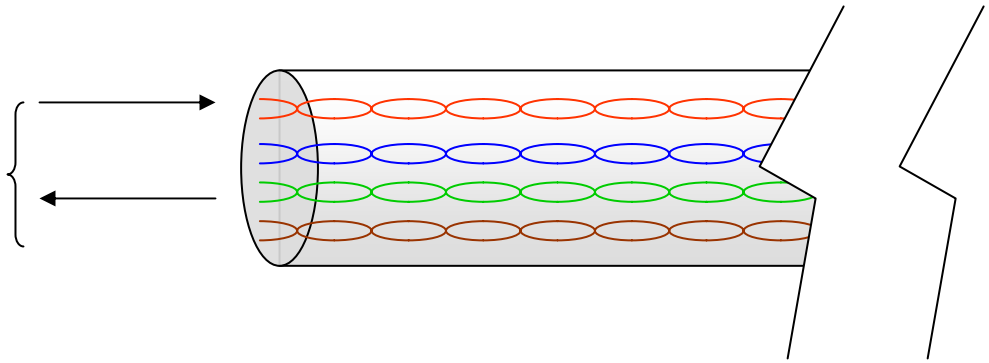
- **Line code: PAM-16 -> PAM-4**
- **Reduced number of channels (each direction): 4 -> 1**
- **Simplex operation at 800MS/S**
  - Produces 1.6Gb/s raw bit rate
- **Zero Stuffing, or equivalent, to match rates**

Note signaling rate remains at 800MS/S

- **Intent is to minimize changes in behavior as an aggressor**

Potential 1G EEE Scheme

Simplex PAM-4 x2
@ 800 Ms/s
= 1.6 Gb/s raw bit rate per direction



Implement simplex PAM-4, which results in a 1.6 Gb/s raw bidirectional rate

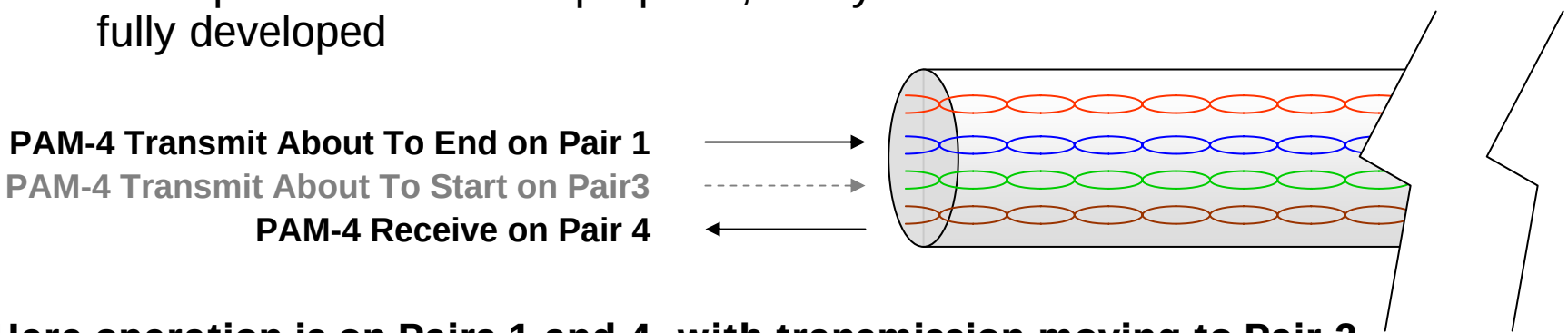
This will allow the echo cancellers, FEXT cancellers, and 11/12 of the NEXT cancellers to be shut off, as well as allow power savings in ¾ the FFEs and THP, the A/D, D/A, hybrid, and transmit power. Error correction will also be much simpler

Maintaining Filters

“Usage Rotation” across all pairs will allow each equalizer to remain updated, along with timing phases

Synchronization of the pair rotation could be coordinated with the 10GBASE-T frame boundary.

- Work is required to determine how much of a training sequence is required to start up a new pair in PAM-4, and how much time is required to switch back to PAM-16 on all four pairs.
- Potentially transition time is less than a single LDPC frame; 320ns
- Concept is the seed of a proposal; many considerations need to be fully developed



Here operation is on Pairs 1 and 4, with transmission moving to Pair 3 once it is stable

LDPC and Framing

Framing in the 1G scheme could use the same frame structure as 10GBASE-T, with zero stuffing to fill in the 10GBASE-T frame

New 1G error correction scheme could use the original LDPC, a modified LDPC, or move to another scheme if desired

- from a power perspective, another error coding scheme may be better, however the PAM-4 LDPC scheme would eliminate the uncoded bits and provide ~12 dB gain over the 10GBASE-T channel

Idle

If extra bits are left in the 10GBASE-T frame after applying the required 1G error correction, the possibility of zero stuffing in the line code exists

For 100M operation, and for further power savings, these gaps could be broadened

Power Estimate

Practical Target: Maintain constant energy per bit

- Should be able to approach 1/5 power for 1/10 data rate for 10GBASE-T operating in 1G subset mode
- Diminishing return when shifting to 1/100 rate

Segment	Power Reduction
A/D	1/4
DAC	1/4
Crosstalk	1/12
Echo	0/4
FFE	1/4
THP	1/4
LDPC	1/x
Tx Power	

Summary

It is possible to achieve fast RPS transitions with a Subset PHY

- 1us may be a practical target

Similar techniques can be applied to 10GBASE-T and to 1000BASE-T

PHY power in 1/10 data rate mode may approach 1/10 power

Implementation complexity potentially less than shifting between “regular” PHYs