100G Backplane PAM4 PHY FEC/PMA Encoding Enhancements

IEEE P802.3bj

May 2012, Minneapolis

Matt Brown – AppliedMicro

Zhongfeng Wang – Broadcom

Contributors and Supporters

- Peter Anslow, Ciena
- Venu Balasubramonian, Cortina
- Stephen Bates, PMC-Sierra
- Sudeep Bhoja, Inphi
- Will Bliss, Broadcom
- Chung-Jue Chen, Broadcom
- Dariush Dabiri, APM
- Arash Farhood, Cortina
- Mark Gustlin, Xilinx
- Ziad Hatab, Vitesse
- Adam Healey, LSI
- Beth Kochuparambil, Cisco
- Kent Lusted, Intel
- Rich Mellitz, Intel
- David Ofelt, Juniper
- Vasu Parthasarathy, Broadcom
- Adee Ran, Intel
- Jamal Riani, Marvell

Introduction

- 100G backplane PAM4 PHY encoding baseline in brown_01a_0312 and brown_02_0312.
- Proposal for
  - RS FEC coding/decoding commonality with NRZ
  - Implementation efficiency
- Results in
  - Improved power efficiency
  - Equivalent/improved latency
  - Equivalent performance
- No changes to NRZ baseline proposal.

Current PAM4 baseline documents:

http://www.ieee802.org/3/bj/public/mar12/brown_01a_0312.pdf

http://www.ieee802.org/3/bj/public/mar12/brown_02_0312.pdf

Considerations regarding current baseline

- PAM4 FEC code word payload is different from NRZ
  - 16 256B/257B blocks per FEC frame, instead of 20 (NRZ).
    - Less commonality and thus less re-use.
  - Rapid alignment marker spacing must be large to be common.
    - For common RAM spacing, RAMs must be spaced by 20 instead of 8, increasing lock time by 2.5x (assuming 8 was desired for NRZ).
- Gearboxing inefficiencies
  - Transcoded payload does not fit exactly into FEC payload so dummy bits and associated gearboxing is required.
  - The 4-bit per-lane PMA overhead insertion causes occasional word rotation.
  - FEC to PMA gearboxing is between two word widths with uncommon factors.

Proposed architectural changes

	Proposal	Baseline
FEC	RS(544,514,T=15,M=10) No dummy bits.	RS(444,412,T=16,M=10) 8 dummy bits per frame
Termination Block	46 bits	64 bits
Overhead	40 bits per 23 FEC frames (~0.12%)	4 bits per 5 FEC frames (~0.07%)
Reference factor	87 (x 156.25 MHz)	88 (x 156.25 MHz)
Symbol rate	13.59375 Gbaud	13.75 Gbaud

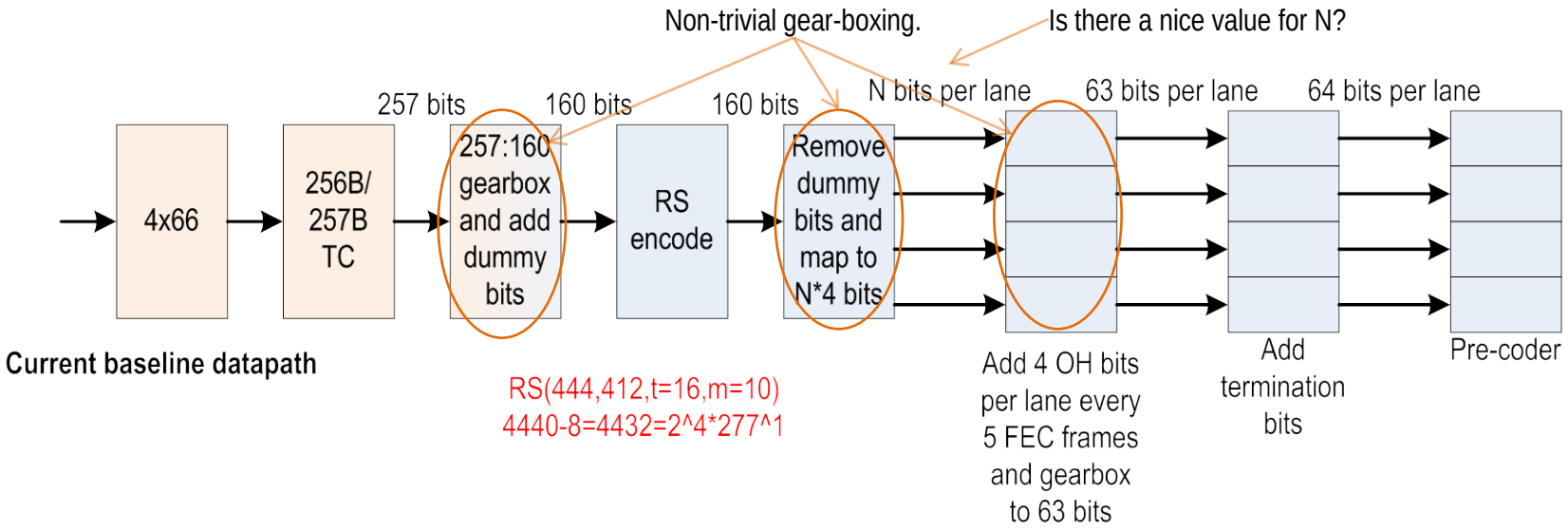
New FEC code word

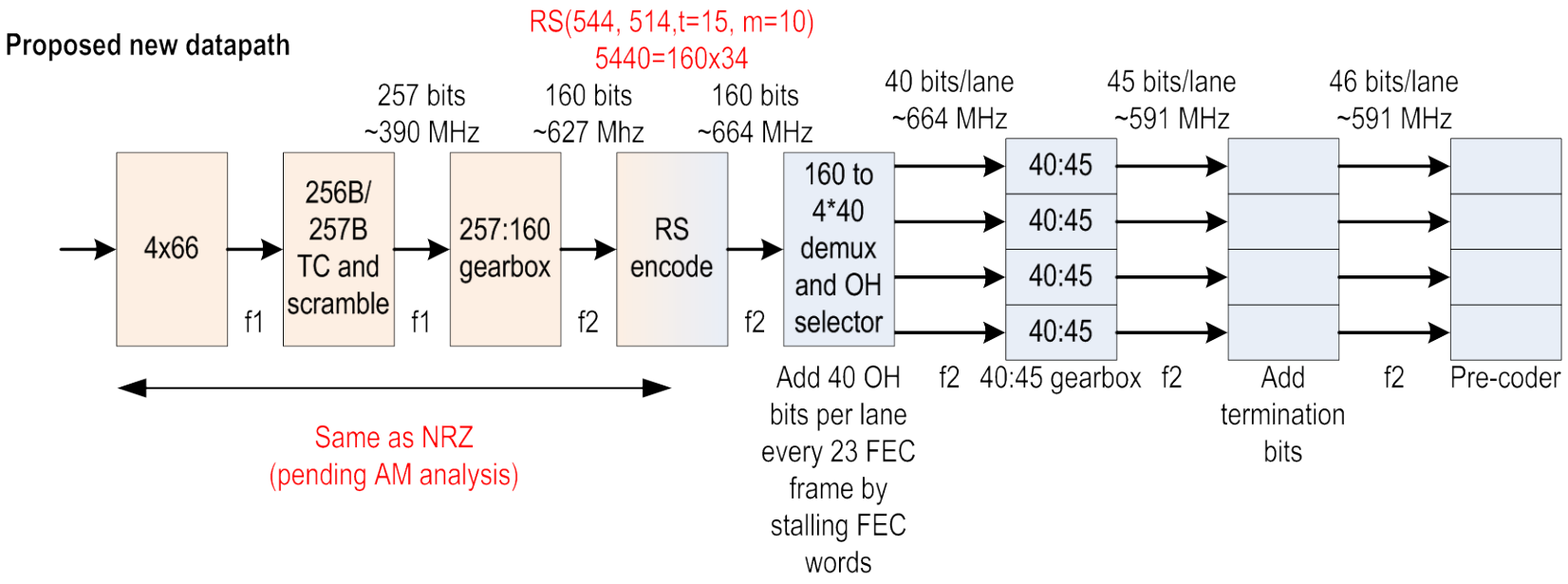
- Use RS(544,514,T=15)
  - Common payload with NRZ.
    - Better opportunity for shared encoding.
    - EEE RAMs fit in the same way.
  - More efficient due to reduced number of corrections and KES time smaller than frame time.
  - Transcoded PCS data fits exactly into payload.
    - No gearboxing for dummy bits.
  - More efficient use of parity.
    - Permits reference clock factor reductions from 88 to 87.
  - Can use the same bus-width for decoder with NRZ
    - 5440/160=34 cycles, 5280/160=33 cycles,
    - Same bus width leads to maximum HW sharing.
  - Net FEC coding gain reduced by less than 0.1 dB.

PMA encoding

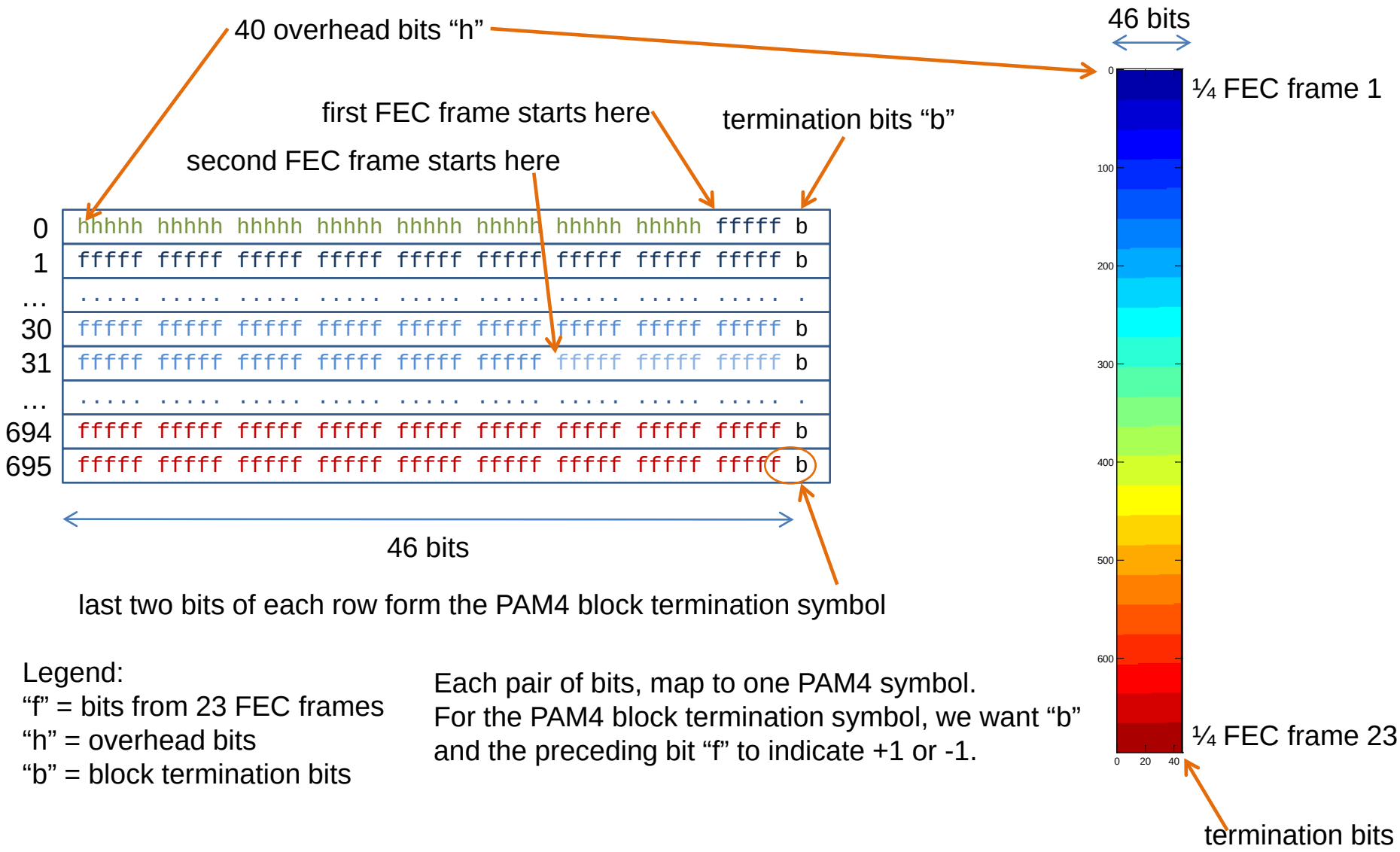
- 46-bit termination block
  - 45 data bits and 1 termination bit per block (9 10-bit FEC words per two termination blocks)
  - Small block size means more efficient, lower latency (~18 ns reduction) in MLSD implementation.
- Overhead 40 bits per 23 FEC frames
  - Using 40-bit bus per lane leads to ~680 MHz transfer rate
  - Requires simple 40:45 bit gearbox to map to termination block
- PMA frame
  - 40 overhead bits
  - 23 * 544 * 10 / 4 = 31280 FEC bits
  - (40+31280) * 1/45 = 696 termination bits
  - Total frame = 40 + 31280 + 696 = 32016 bits
- PMA frame period = 1.18 us
  - If using overhead word for sync, still possible to lock much faster than using AM.
    - In EEE rapid alignments marker or alternate are required for fast wake time.

Gearboxing example





PMA frame structure (one per lane)



Legend:
“f” = bits from 23 FEC frames
“h” = overhead bits
“b” = block termination bits

Each pair of bits, map to one PAM4 symbol.
For the PAM4 block termination symbol, we want “b” and the preceding bit “f” to indicate +1 or -1.

Latency analysis

- The incremental latency budget is shown in the table below.
- See details in backup slides.

Item	Latency savings
Larger FEC frame size.	-10 ns
For FEC, decoding t=15 instead of t=16.	+2.5 ns
Gearbox simplifications.	+7.5 ns
More efficient MLSD RX.	+18 ns
Net savings.	+18 ns (0 ns if no RX improvement)

Open issues

- This presentation does not deal with the following open items
  - Synchronization methodology for normal or EEE operation.
  - Scrambling (or not) of alignment markers.
    - See anslow_01_0512.
- More study is required to determine the optimal arrangement of the termination bits and OH.

Summary

- Modify baseline as follows:
  - Change FEC code word to RS(544,514,t=15,m=10).
  - Map FEC payload the same as for NRZ. No dummy bits required.
  - Change PMA termination block size to 46 bits.
  - Change PMA frame overhead to 40 bits every 23 FEC frames.
  - Change PAM4 symbol rate to 13.59375 GHz (87 * 156.25 MHz)
- Modifications result in:
  - Common (with NRZ) 256B/257B block to FEC mapping.
  - Common (with NRZ) rapid alignment marker spacing
  - Significant implementations efficiencies.
  - Latency reduction between 0 and 18 ns depending on implementation.
  - 1.1% symbol rate reduction.

Conclusion

- Straight forward modifications to the baseline FEC and PMA encoding.
- Encoding commonality with NRZ.
- Implementation efficiency.
- No compromise in latency, power, or performance.

Thanks!

Backup slides

Coding parameters considered

FEC	Ref Mult. Factor	OH (bits)	OH space (FEC blocks)	OH period	T space (bits)	Coding penalty (dB)	Pro	Con
t=15	87	40	23	1.18 us	46	<0.1	40b OH Optimal gearbox	OH period
t=15	86-2/3	None	N/A	N/A	52	~0	No OH Lower symbol rate	Fractional PLL 40:51 gearbox
t=15	87	50	7	358 ns	56	<0.1		50b OH 50:55 gearbox
t=15	87	40	5	256 ns	58	<0.1	40b OH	40:57 gearbox
t=15	87	32	4	205 ns	58	<0.1	Short OH period.	32:57 gearbox
t=15	87	40	110	5.6 us	44	<0.1	40b OH	OH not useful
t=15	87.5	None	N/A	N/A	36	~0.2	No OH	Fractional PLL
t=16	87.5	None	N/A	N/A	40	~0	No OH	Fractional PLL
t=15	88	40	10	512 ns	32	>0.2	40b OH	Coding gain penalty
t=16	88	80	5	256 ns	44	<0.1	80b OH	Coding gain penalty

Coding gain comparison

Code	Total Coding gain (dB)	Burst channel coding gain (dB)	Over Clocking Loss (dB)	Avail. Gain (dB)
Current Baseline RS(444, 412, 16, m=10)	7.34	6.51	(6.7%) 0.99	5.52
This proposal RS(544, 514, 15, m=10)	7.10	6.26	(5.5%) 0.82	5.44

Latency change estimation

- Assume the system clock period is 1.25ns (i.e., 800Mhz), which is aggressive.
- The computation time of KES in RS decoding is 2*t cycles. For t=15 compared to t=16, we save 2 cycles (2.5 ns).
- Regarding gearbox, the new proposal save 3 gearbox operations in TX and RX. In total, we save 6 cycles (7.5 ns).
- If using MLSD, we may need 27 parallel VD working in parallel to achieve 27Gbps throughput, where we assume each VD works at about 500Mhz. Thus the saved latency with MLSD is 27*(64b-46b)/27G =18ns.