Data Rate to Line Rate Conversion

Glen Kramer (Broadcom Ltd)

Motivation

- MAC data rate is 25 Gb/s
- 25GMII transmits 32 bits @ 390.625 MHz (on both rising and falling edges)
- 64b/66b encoder adds 3.125% of overhead and correspondingly increases the transmission rate to 25.78125 Gb/s (25x66/64)

- Now we have an emerging consensus to use 256b/257b line coding.
  - This scheme reduces the overhead from 3.125% to 0.390625%.
  - If we keep the line rate the same, the new and improved throughput is 25.78125 * 256/257 = 25.68093385214007782101167315175 1…Gb/s
  - Where does the extra data come from to fill the increased throughput capacity?
  - How do we document this irrational data rate?

Two approaches

Two general approaches:

1. We can speed-up MAC/MPRS/25GMII data path to fill the increased capacity
   - This method is assumed in kramer_3ca_1_0118.pdf

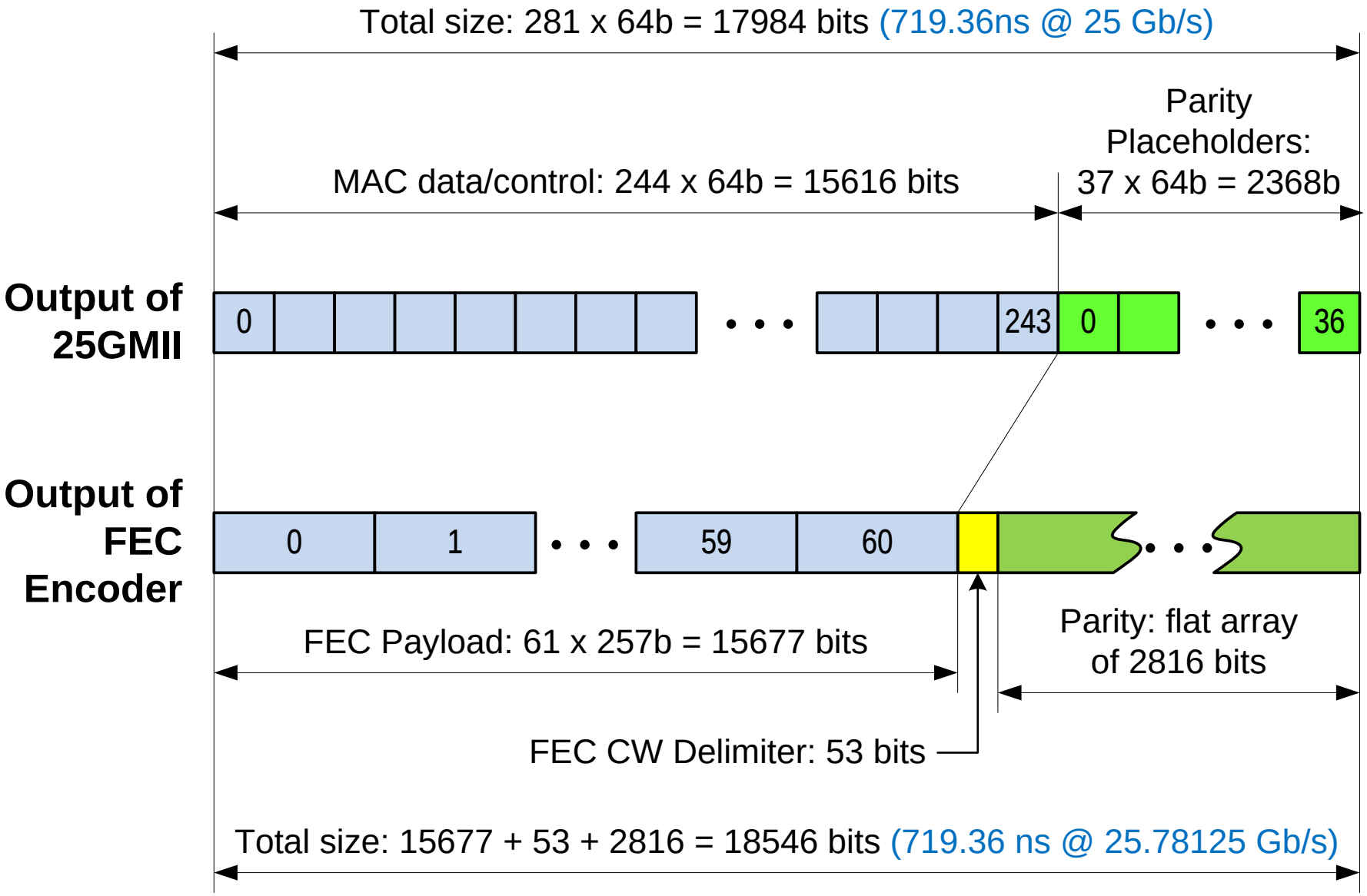
2. We can keep MAC/MPRS/25GMII data rate as is and inflate the data below 25GMII to fill the extra capacity
   - This method is outlined in gao_3ca_1_0118.pdf
   - For example, MPRS would generate fewer parity placeholders than needed for the FEC encoding.
   - PCS/FEC encoder adds the actual parity. This inflates the data to fill the extra capacity

- Any other method?

2nd approach

- 1 full FEC codeword takes 72x257b @ 25.78125 Gb/s = 717.7309(09) ns
- 25GMII transmits 64 bits every 2.56 ns, so in 717.7309(09) ns we only receive 280.36(36) blocks. But we need 72x4=288 64b blocks to form the FEC codeword!
- We can have 8 extra blocks added by FEC encoder, but what to do with extra 0.363636… of a 64b block?
- The only solution here is to extend the FEC codeword to the same time that 281 64b blocks take through the 25GMII (281 x 2.56 = 719.36 ns)
- This will require FEC codeword to have 42 bits of extra padding that are outside of any 256b/257b block
- The FEC+encoding efficiency in this case is (61 x 256b)/18546 = 84.2%, which is still ok.

Illustration of 2nd approach



1st approach

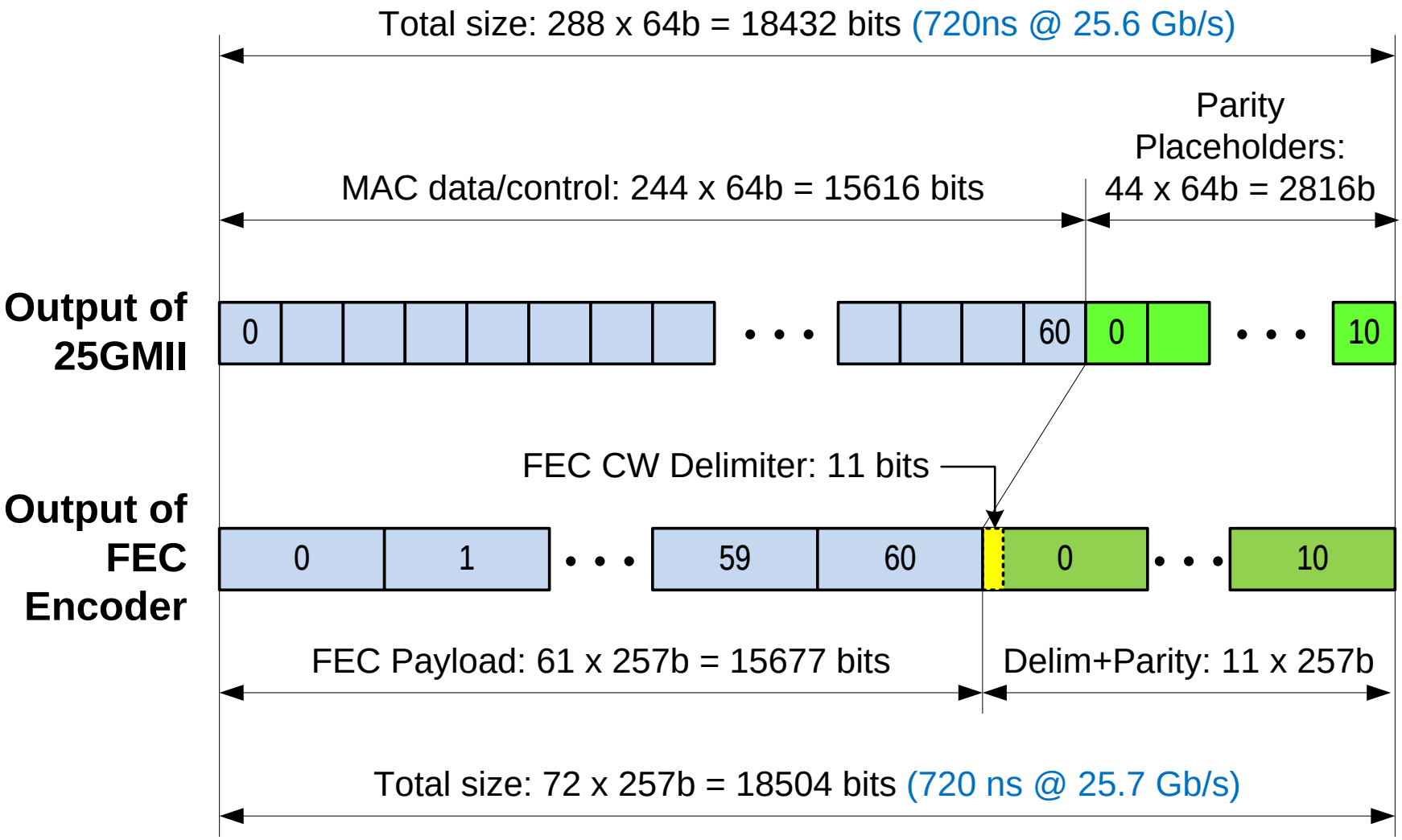
- This approach keeps the MPRS SDs as they are, just changes the 25GMII clock. Since MPRS is driven by the 25GMII clock, and MAC is driven by the MPRS clock, we only need to redefine the 25GMII clock for the data to fully utilize the increased capacity.
- It also would be nice if the rate of 257b blocks on the line is some nice rational number.

- The 802.3by specified 25GMII simply by reference to Cl. 49:

> **106.3 25GMII functional specifications**
>
> The 25GMII functions identically to the XGMII specified in 46.3 with the exception that the TX_CLK and RX_CLK frequency shall be 390.625 MHz ±100 ppm (also one-sixty-fourth of the MAC transmit data rate).

- We can do the same, but specify different RX_CLK and TX_CLK.
- What clock would be good?

Illustration of 1st approach



25GMII and line rates are shown as examples. Rates are discussed on the next slide.

What is a good clock?

	25GMII Clock ±100ppm (MHz)				
	390.625	**400**	**401**	**401.25**	**401.2645914**
MAC data rate (Gb/s)	25	25.6	25.664	25.68	25.6809338..
EQ time (ns)	2.56	2.5	2.493765586	2.492211838	2.4921(21)
FEC codeword time (ns)	737.28	720	718.2044888	717.7570093	717.7309(09)
Line rate (Gb/s)	25.09765625	25.7	25.76425	25.7803125	25.78125
Throughput Delta (Mb/s) (1st approach – 2nd approach)	-529	-20	34	47	48

Nice numbers

Line Rate = 25.78125 Gb/s

Timing values per 2nd approach

25GMII Clock (MHz)	390.625 ±100ppm
MAC data rate (Gb/s)	25
EQ time (ns)	2.56
FEC codeword time (ns)	719.36
Line rate (Gb/s)	25.78125

Side-by-side comparison

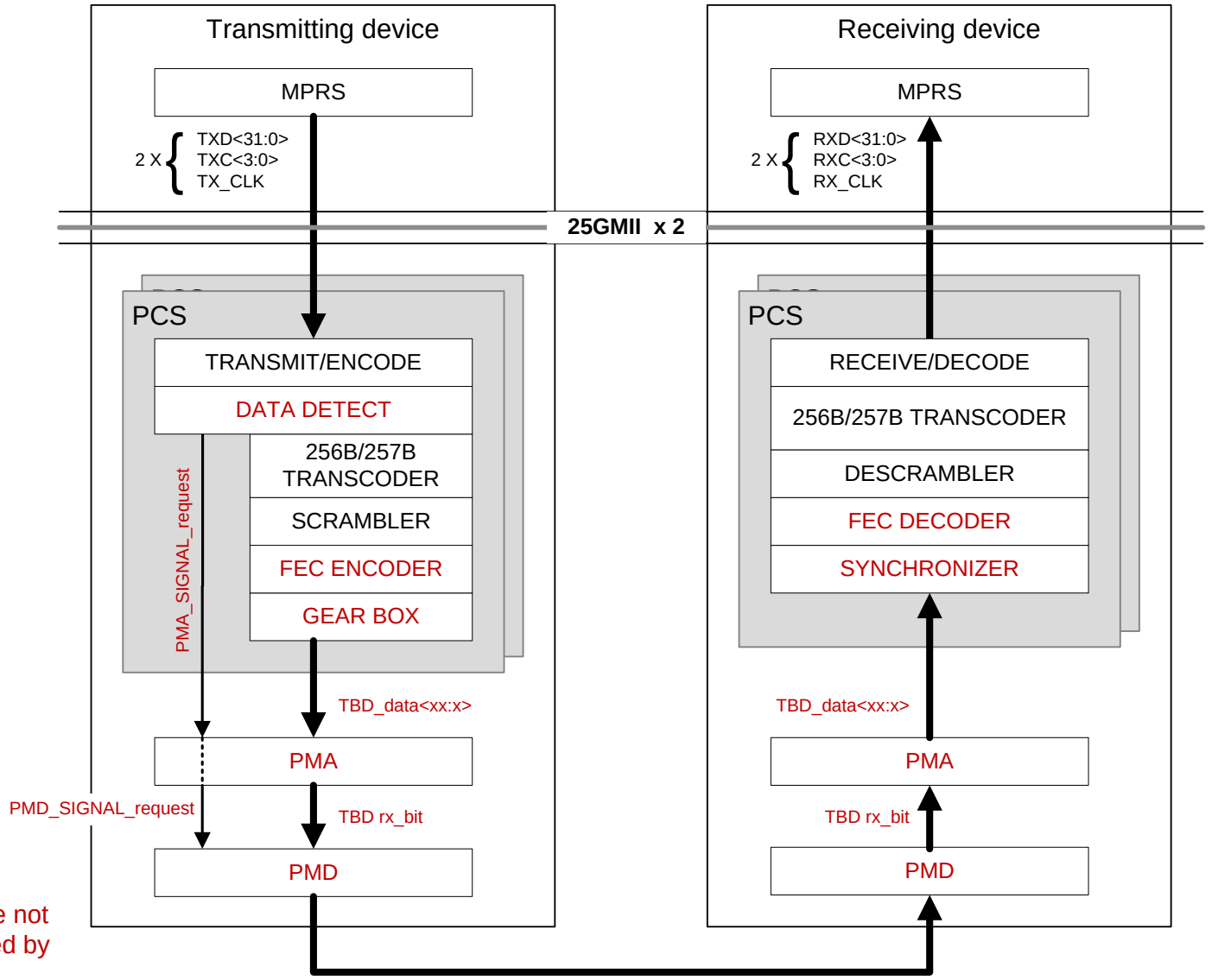
Approach #1

- Treats 256b/257b line coding as primary. FEC CW consists of 72 257b blocks.
- Approach #1 gives us a chance to unify line rate with ITU-T 25G-PON (too speculative at this time)
- Slightly better FEC information rate (84.4% vs. 84.2%)

Approach #2

- Treats FEC framing as primary. The 256b/257b coding is applied only within the FEC payload.
- Gives us extra 42 bits (53 total) to delineate FEC codewords (is useful for the downstream)

Location of State Diagram



Blocks in red have not yet been discussed by the TF

Motion #3

- Accept 256B/257B line coding for downstream

- Technical (>= 75%)
- Moved: Marek Hajduczenia
- Seconded: Glen Kramer

- Y:21 N:0 A:4
- Motion passed

Motion #6

- Accept the rate adjustment mechanism as outlined in the 2nd approach on slides 4-5 of kramer_3a_4a_0118.pdf (based on gao_3ca_1_0118.pdf, slide 4).

- Technical (>= 75%)
- Moved: Glen Kramer
- Seconded: Gao Bo

- Y:25 N:0 A:5