

10GE WAN PHY: Physical Coding Sublayer (PCS)

**IEEE 802.3 HSSG Meeting, Dallas
January 18-20, 2000**

**Norival Figueira, Paul Bottorff – Nortel Networks
Pankaj Kumar – Level One/Intel
Tom Palkert – AMCC
Iain Verigin – PMC-Sierra
Nader Vjeh – Lantern Communications
Frederick Weniger – Vitesse**

Based on Document

- **“Proposal for a 10 Gigabit Ethernet WAN PHY”**

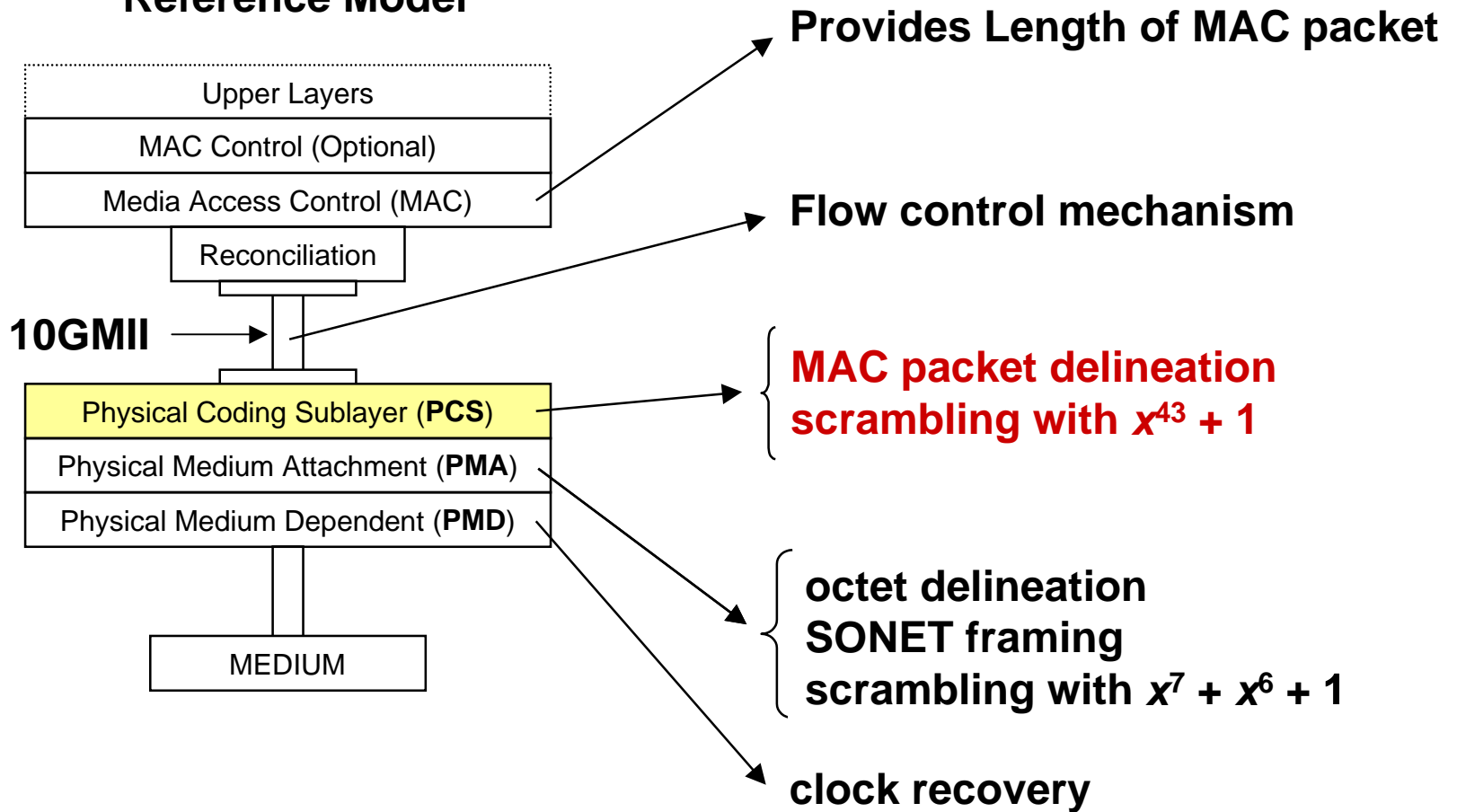
— http://grouper.ieee.org/groups/802/3/10G_study/public/nov99/figueira_2_1199.pdf

Agenda

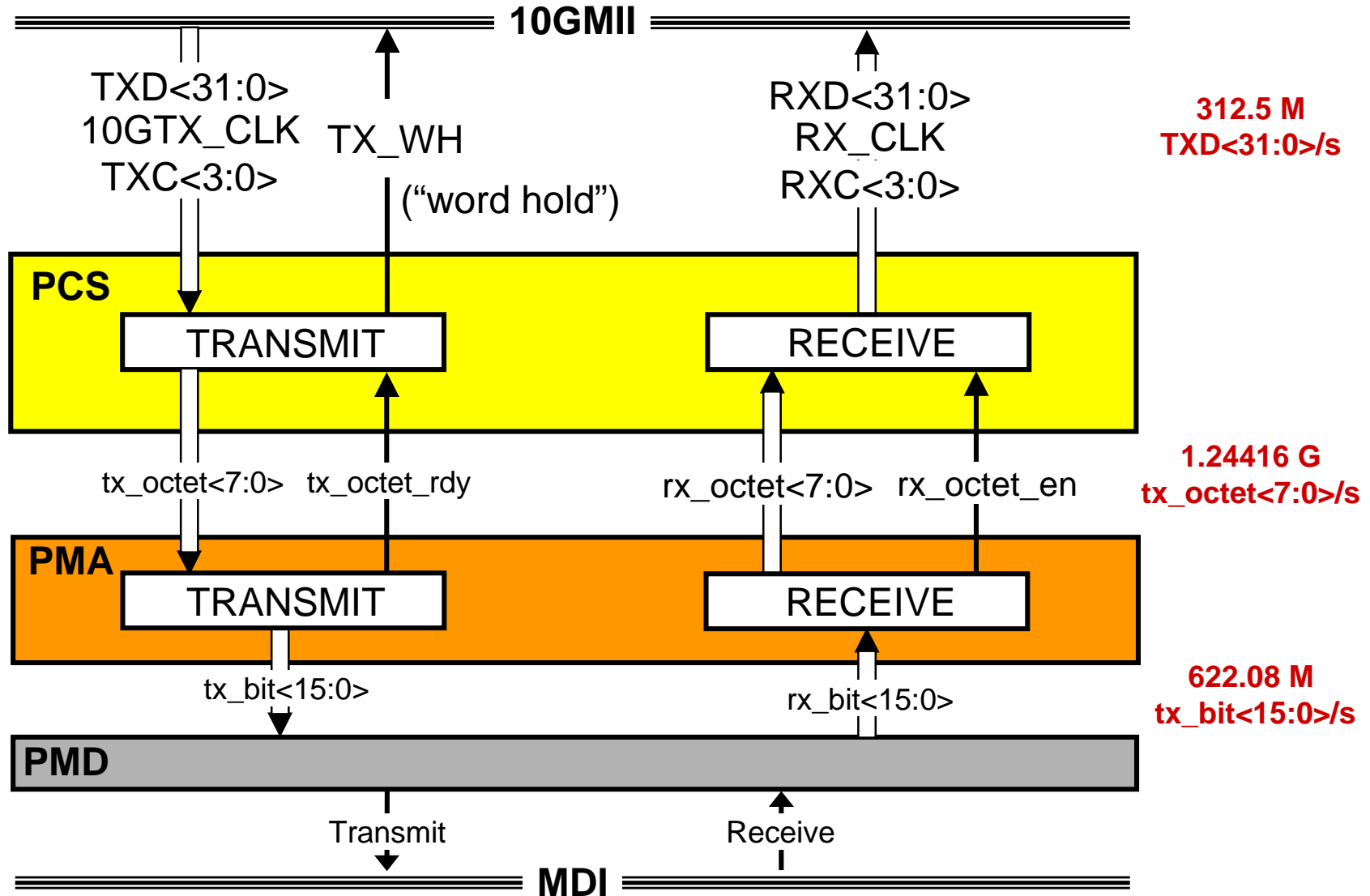
- **10GMII**
- **PCS Encapsulation**
 - octet stream at 10GMII and PCS-PMA interface
- **PCS Transmit and Receive processes**
- **MAC packet delineation**
 - use of header error control (HEC) check algorithm
 - Idle PHY packets
 - state diagrams
- **$x^{43}+1$ self-synchronous scrambler**

Architectural Positioning

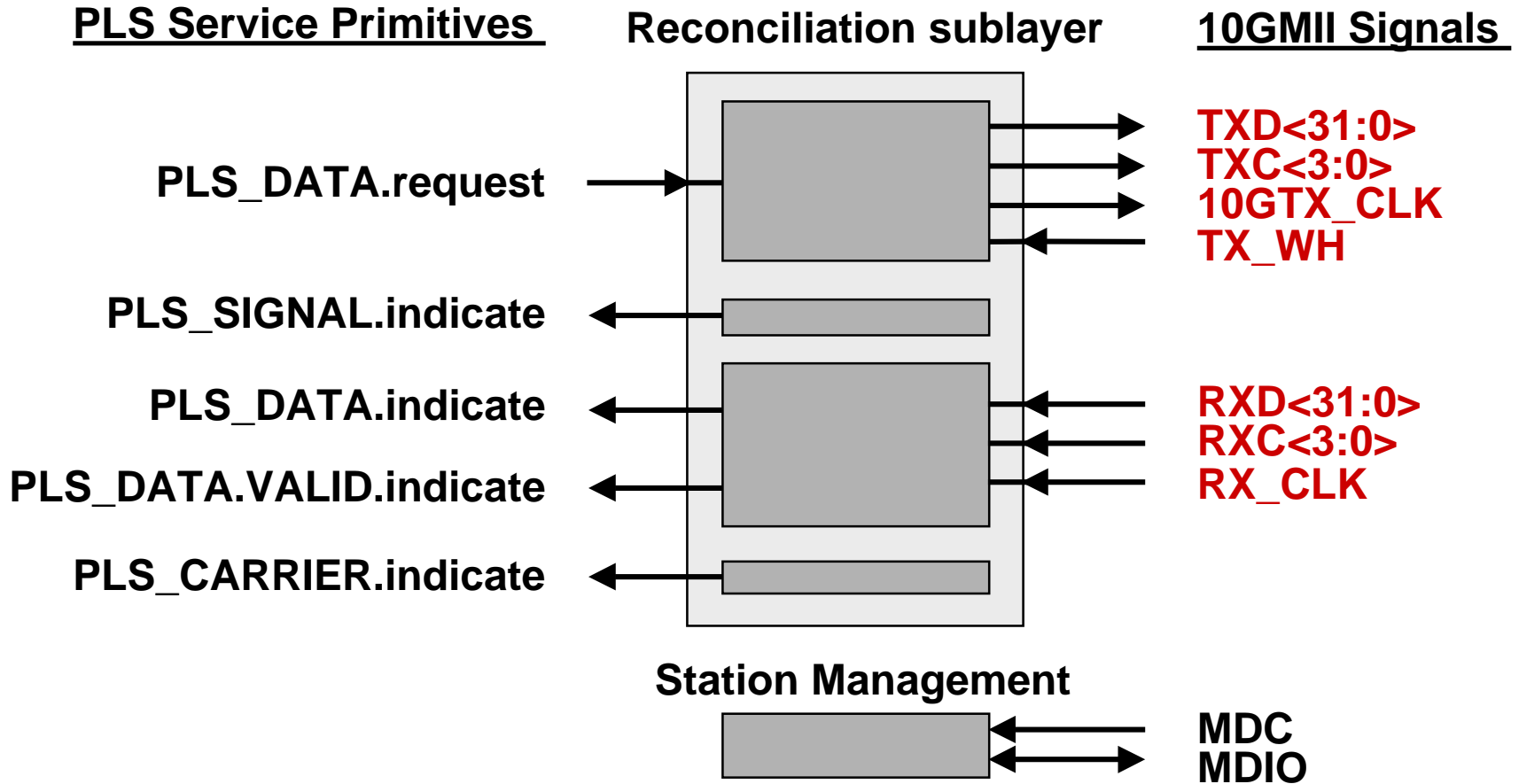
10 Gigabit Ethernet Reference Model



Functional Block Diagram



Reconciliation Sublayer and 10GMII



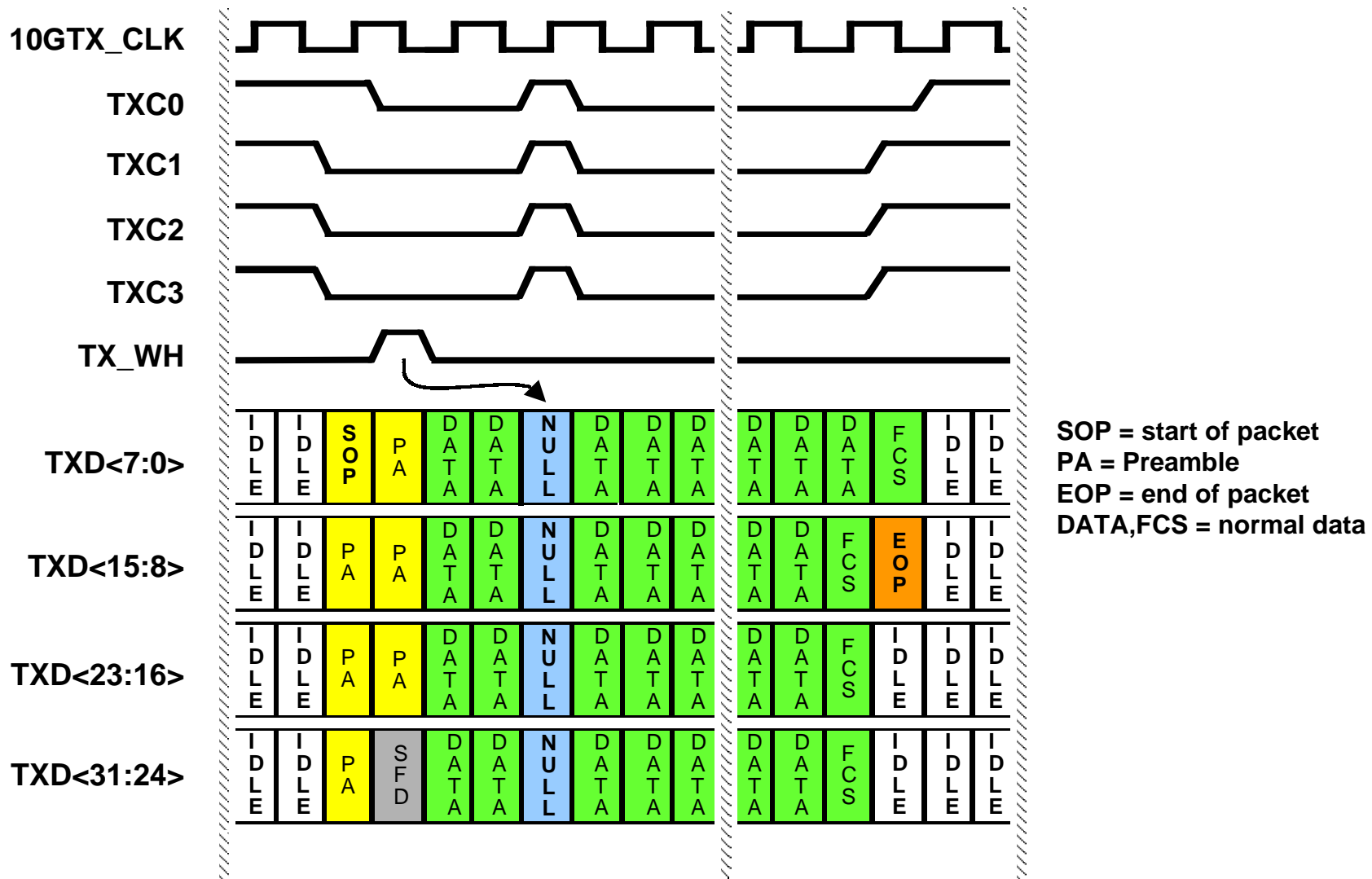
10GMII

- **Builds on Frazier et al MAC/PCS parallel interface**
- **Uses embedded special control signals**
 - TXC<3:0> and RXC<3:0>
 - When asserted, a control symbol is conveyed on the data path
 - Control symbols: SOP, IDLE, NULL, or EOP
 - When de-asserted, data is conveyed on the data path
- **Provides 4-byte wide data path**
 - RXC<31:0> and TXC<31:0>

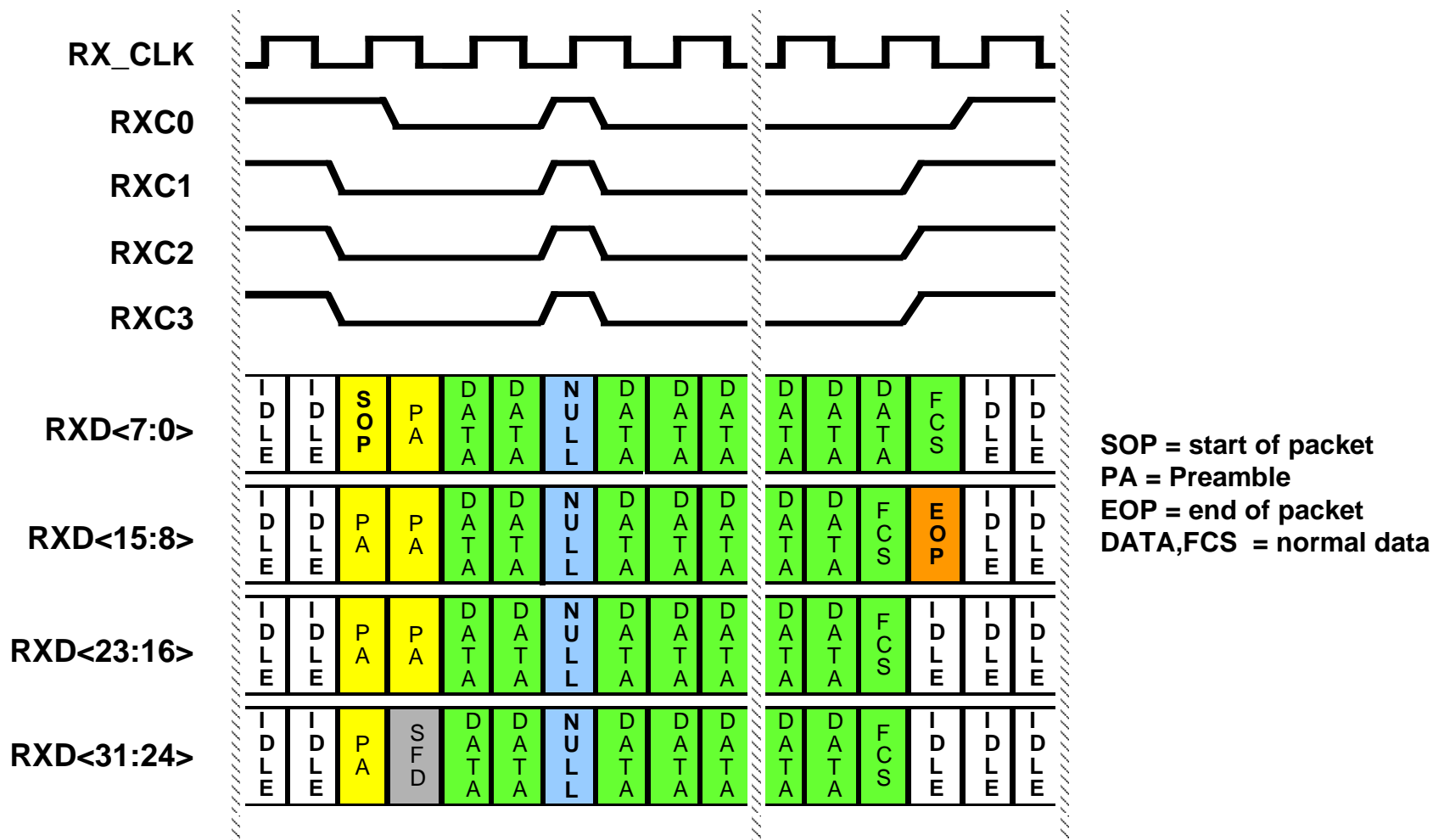
10GMII (cont.)

- **Uses dual data rate (DDR) signaling at 156.25 MHz**
 - Data and control signals are sampled on both rising and falling edges of the clock
 - Maximum effective data rate of 10 Gb/s
- **NULL character is used with TX_WH signal to provide a word-based flow control mechanism**
 - Effective data rate is flow controlled to 9.58464 Gb/s

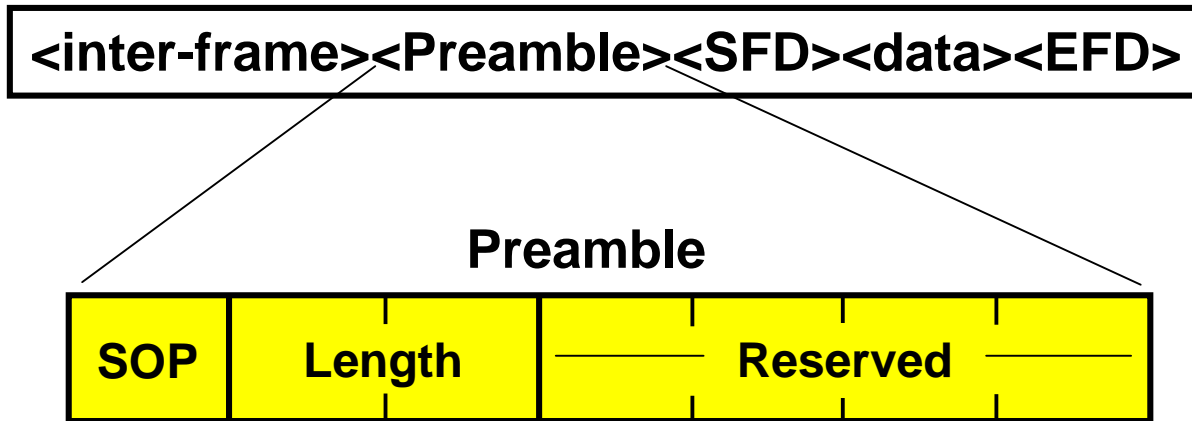
10GMII: Basic Frame Transmission



10GMII: Basic Frame Reception



10GMII: Data Stream



- **Length**

- number of octets of the MAC packet, i.e., length of <data> (DA to last octet of FCS field)

- **Reserved**

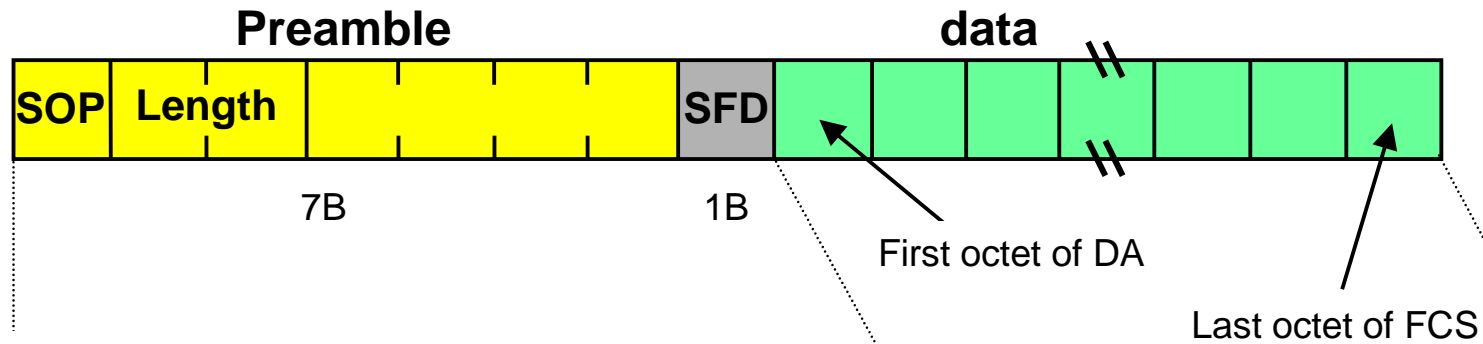
- can be used for: time to live, MPLS label, congestion notification, per-hop behavior

Functions within the PCS

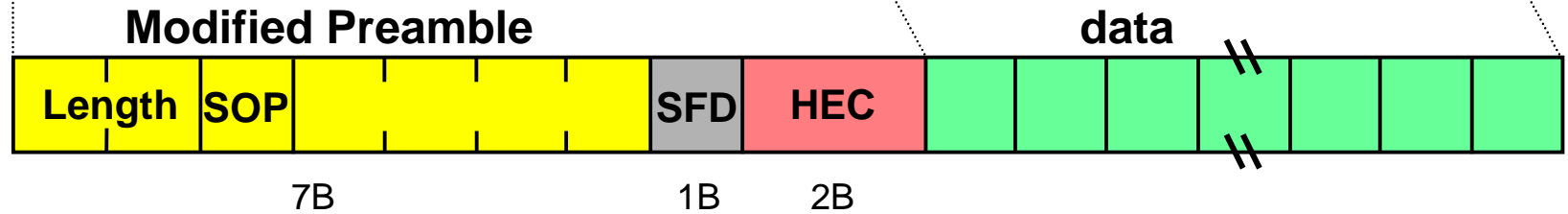
- **Scrambling and descrambling of data octets**
 - $x^{43} + 1$ self-synchronous scrambler
- **Packet delineation using the Header Error Control (HEC) check algorithm**
- **Communication with underlying PMA sublayer**
- **Communication with the Reconciliation sublayer**

PCS Encapsulation

Octet stream at 10GMII



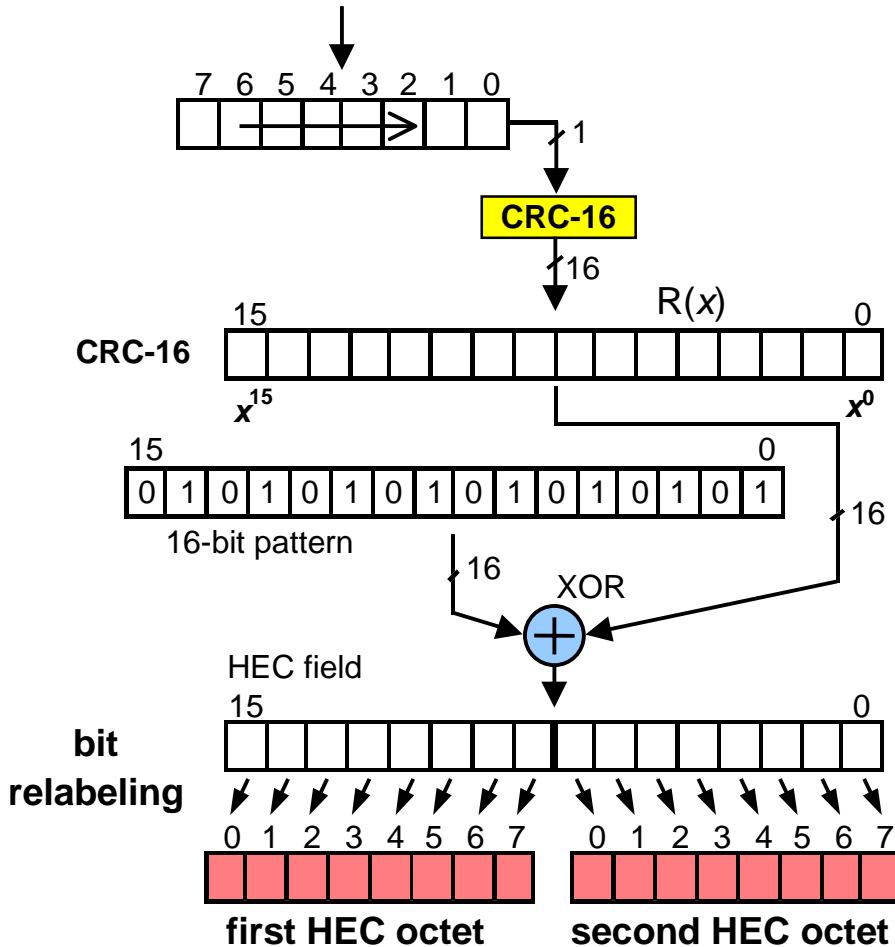
Octet stream at PCS-PMA I/F



HEC (header error control) = CRC-16 of Modified Preamble and SFD

HEC Calculation

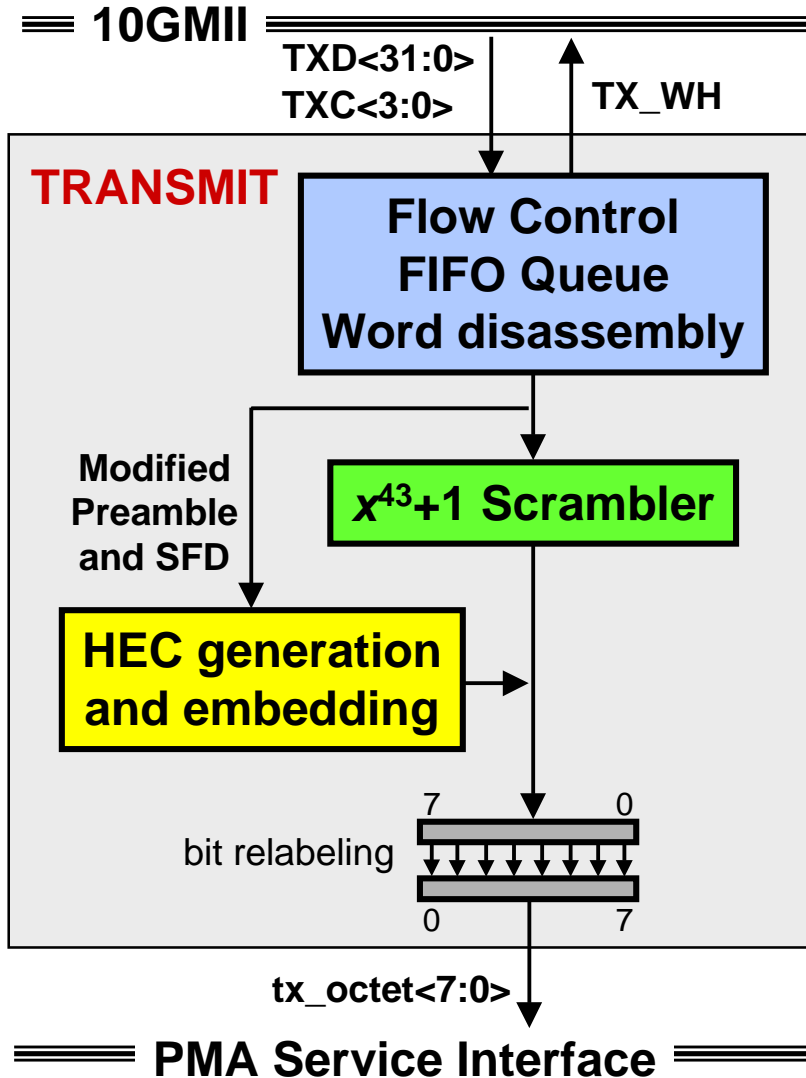
(8 octets of Modified Preamble + SFD in order)



(functional diagram)

- **CRC-16 calculated LSB first**
- **$G(x) = x^{16} + x^{12} + x^5 + 1$**
- **Remainder is added module 2 to 0101010101010101 to improve packet delineation**
- **Bits are relabeled to agree with Ethernet bit transmission order**

PCS Reference Diagram: Transmit



- **Flow Control Unit**

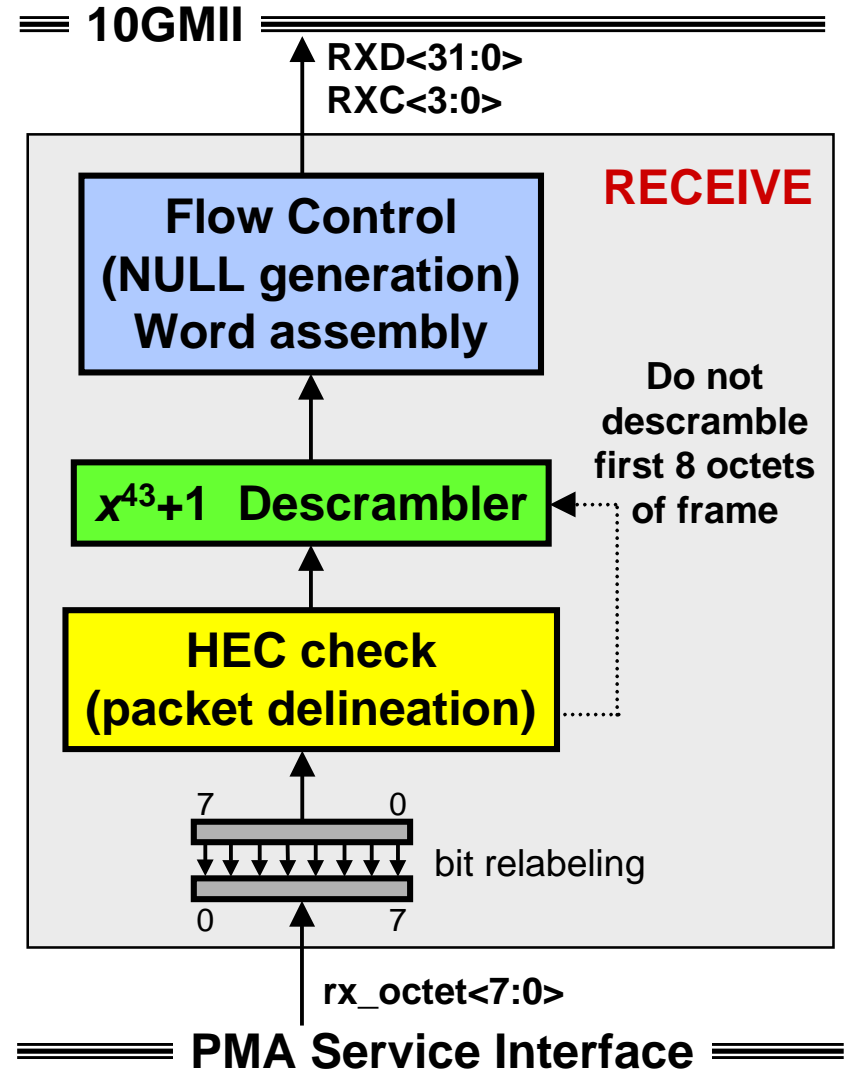
- Provides rate match from 10 Gb/s to 9.95328 Gb/s
- Allows flow control from PMA (for overhead embedding)
- Requests NULL words with TX_WH signal to avoid FIFO overflow

- **Bit relabeling required to maintain usual FCS error detection capabilities**

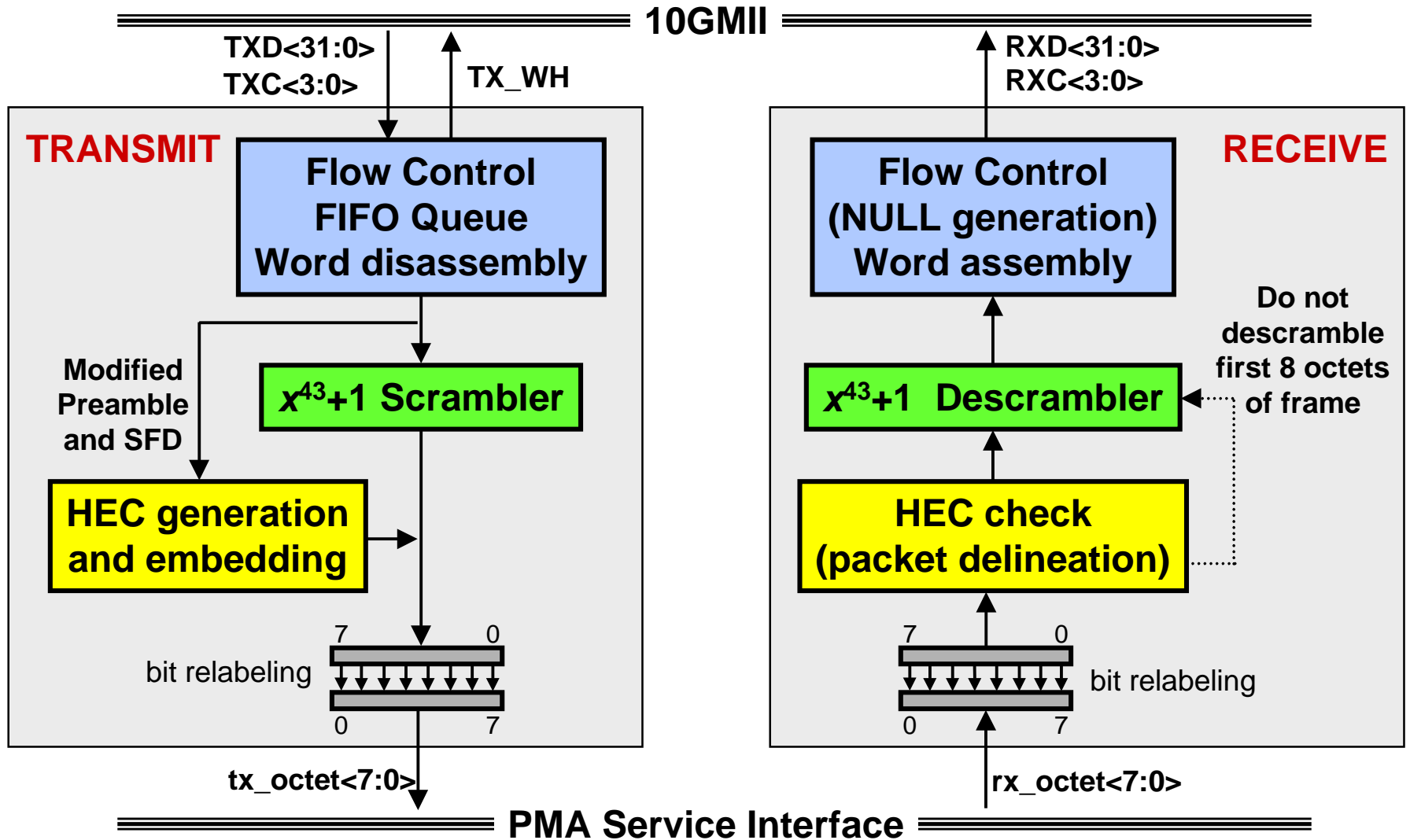
- Because PMD transmits MSB first
- No change to burst errors

PCS Reference Diagram: Receive

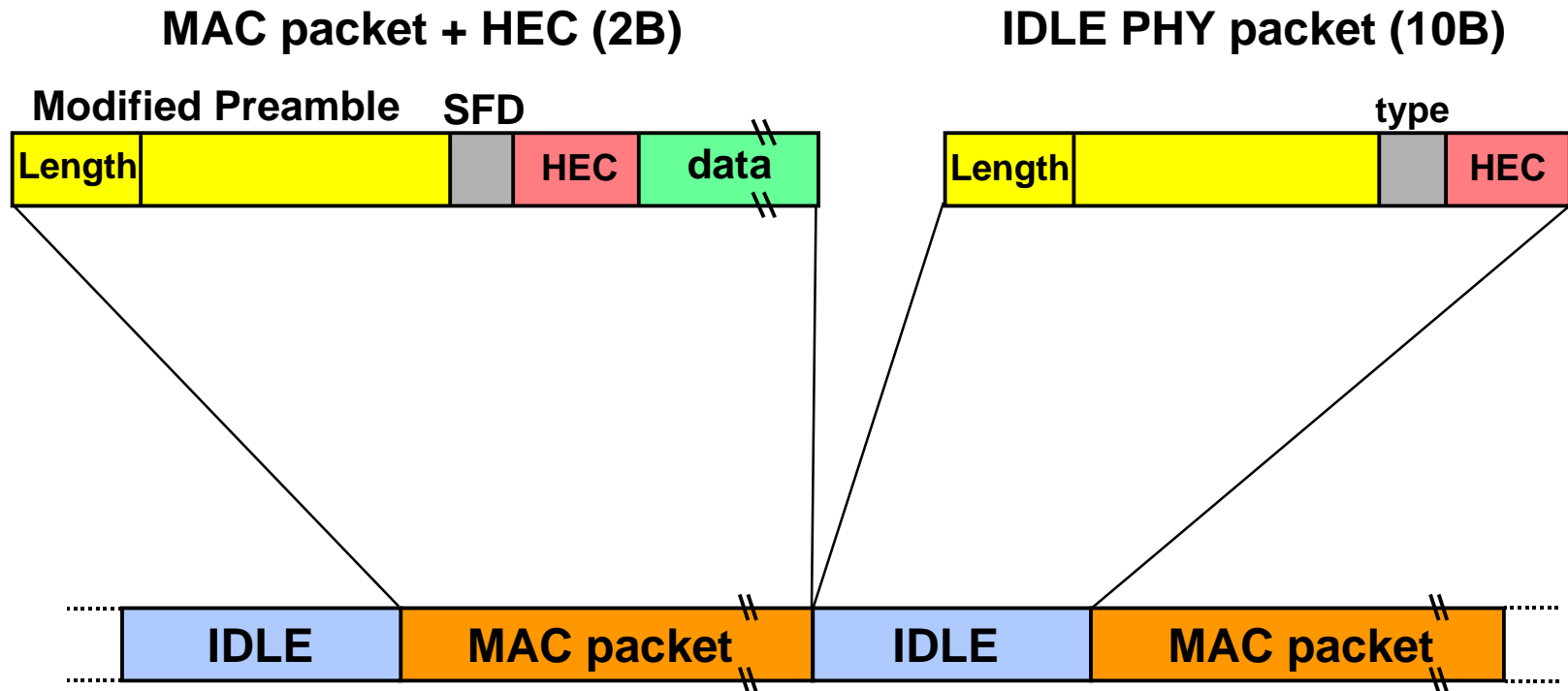
- **Flow Control Unit**
 - Provides rate match from 9.95328 Gb/s to 10 Gb/s
 - Allows flow control from PMA (to skip overheads)
- **Word assembly provides alignment to word boundaries**



PCS Reference Diagram

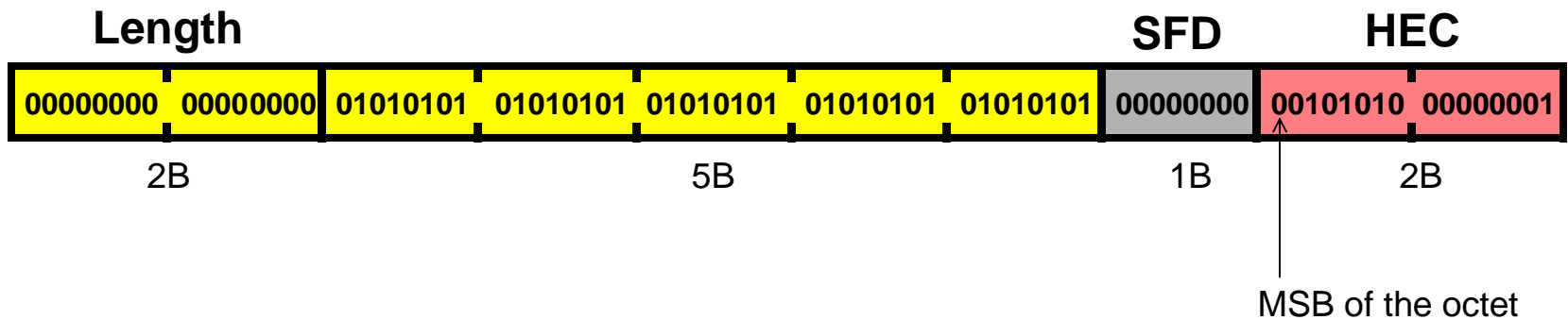


PCS Encapsulation (cont.)



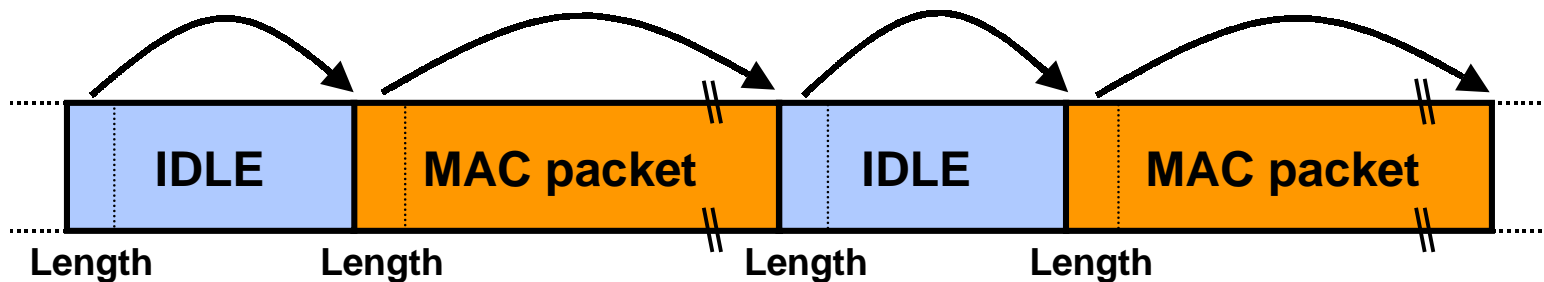
Idle PHY Packet

- Generated whenever the PCS Transmit process needs to generate an octet and there is no MAC packet available for transmission
- Cause no action at PCS Receive process except for packet delineation including HEC verification



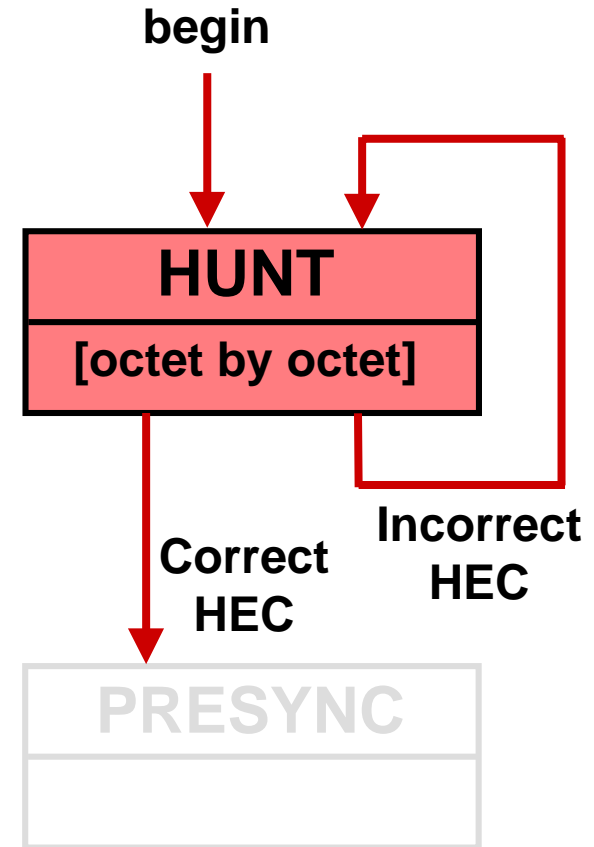
Packet Delineation

- Modified version of the **HEC check algorithm** specified in ITU- I.432
- Based on the correlation between the (modified) Preamble/SFD and the embedded HEC field
- Length field is used to find the beginning of the next packet

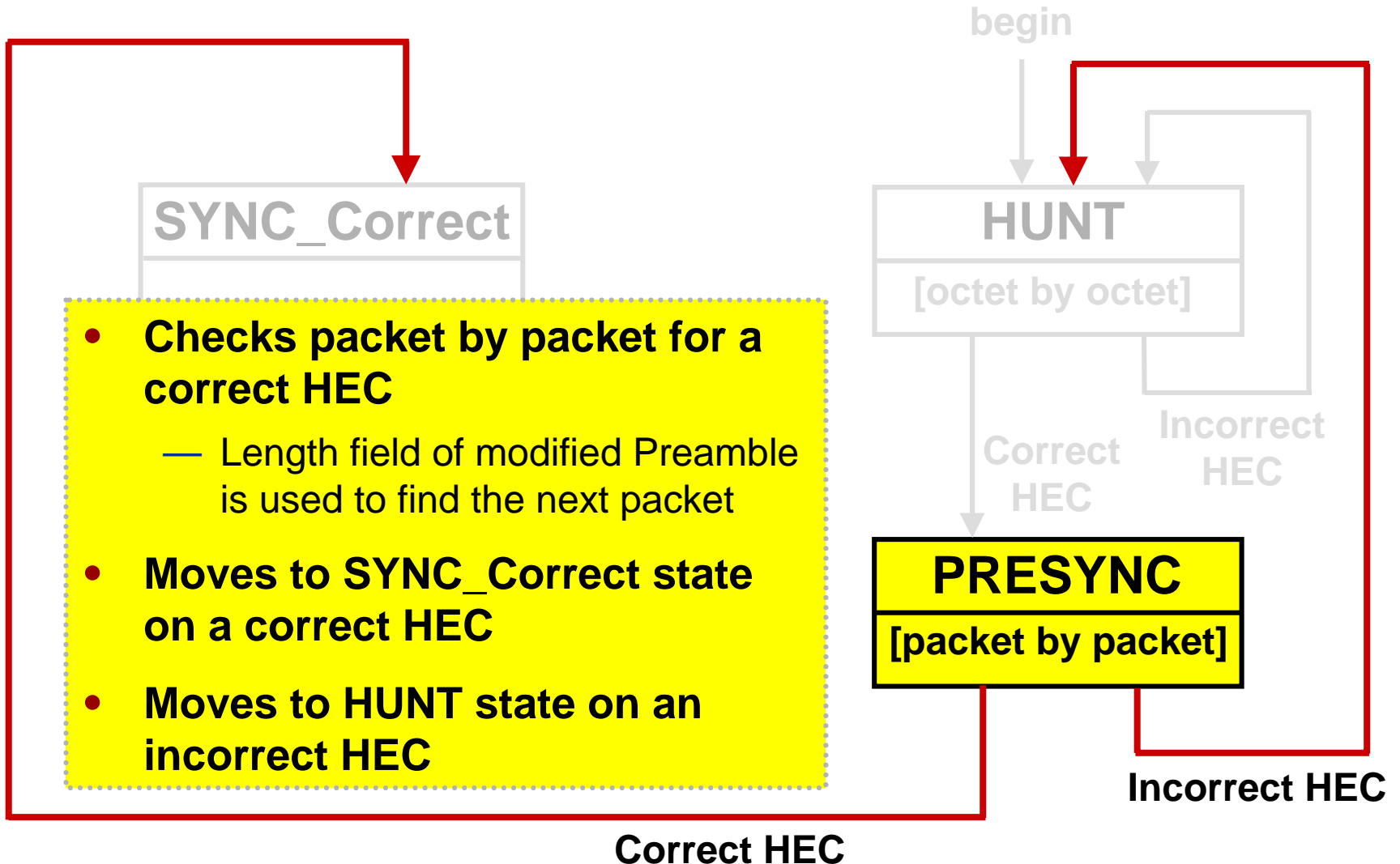


Packet Delineation: HUNT state

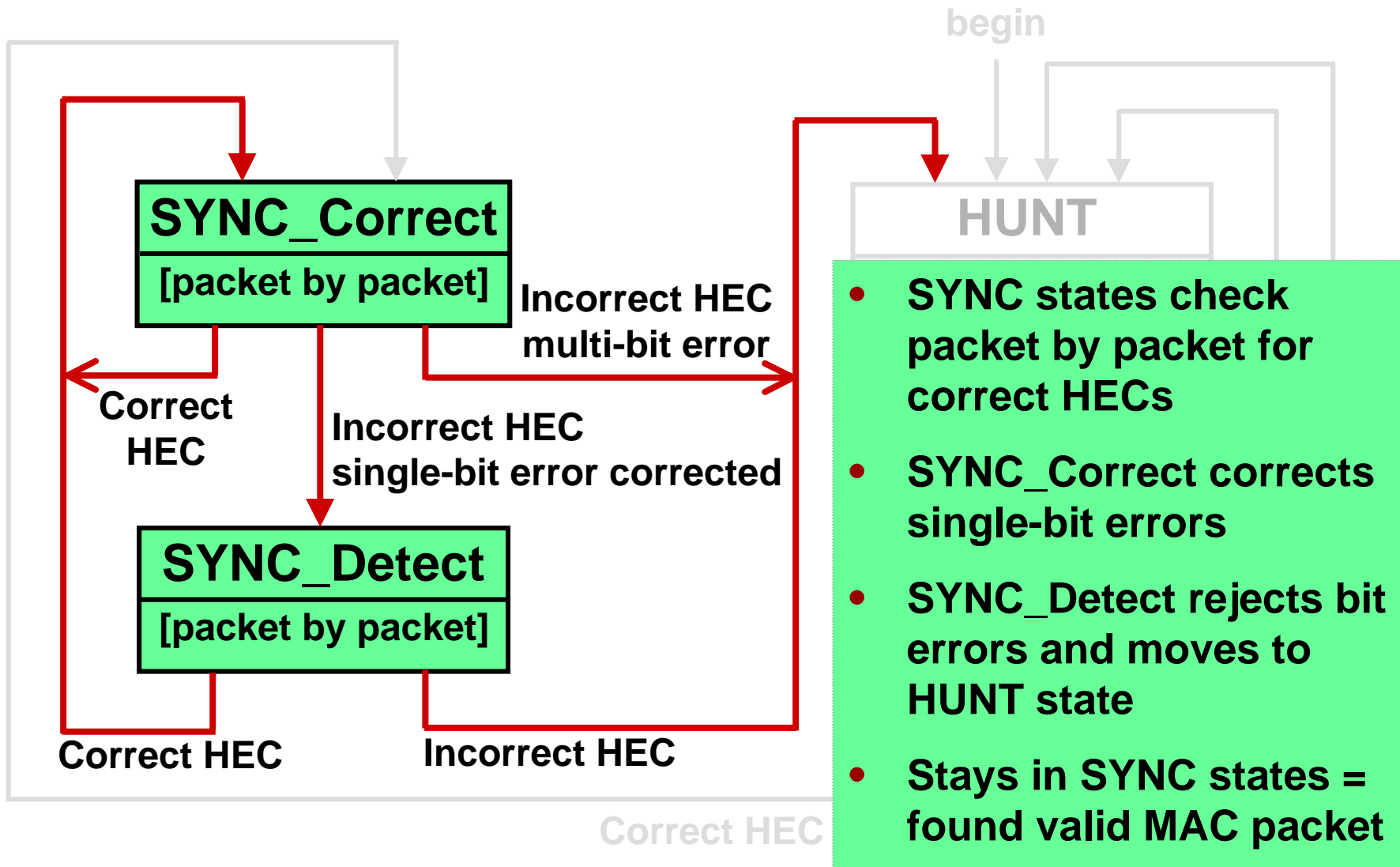
- **Initial state**
- **Checks octet by octet for a correct HEC**
 - Octet delineation is done at the PMA sublayer
- **Moves to PRESYNC state on a correct HEC**



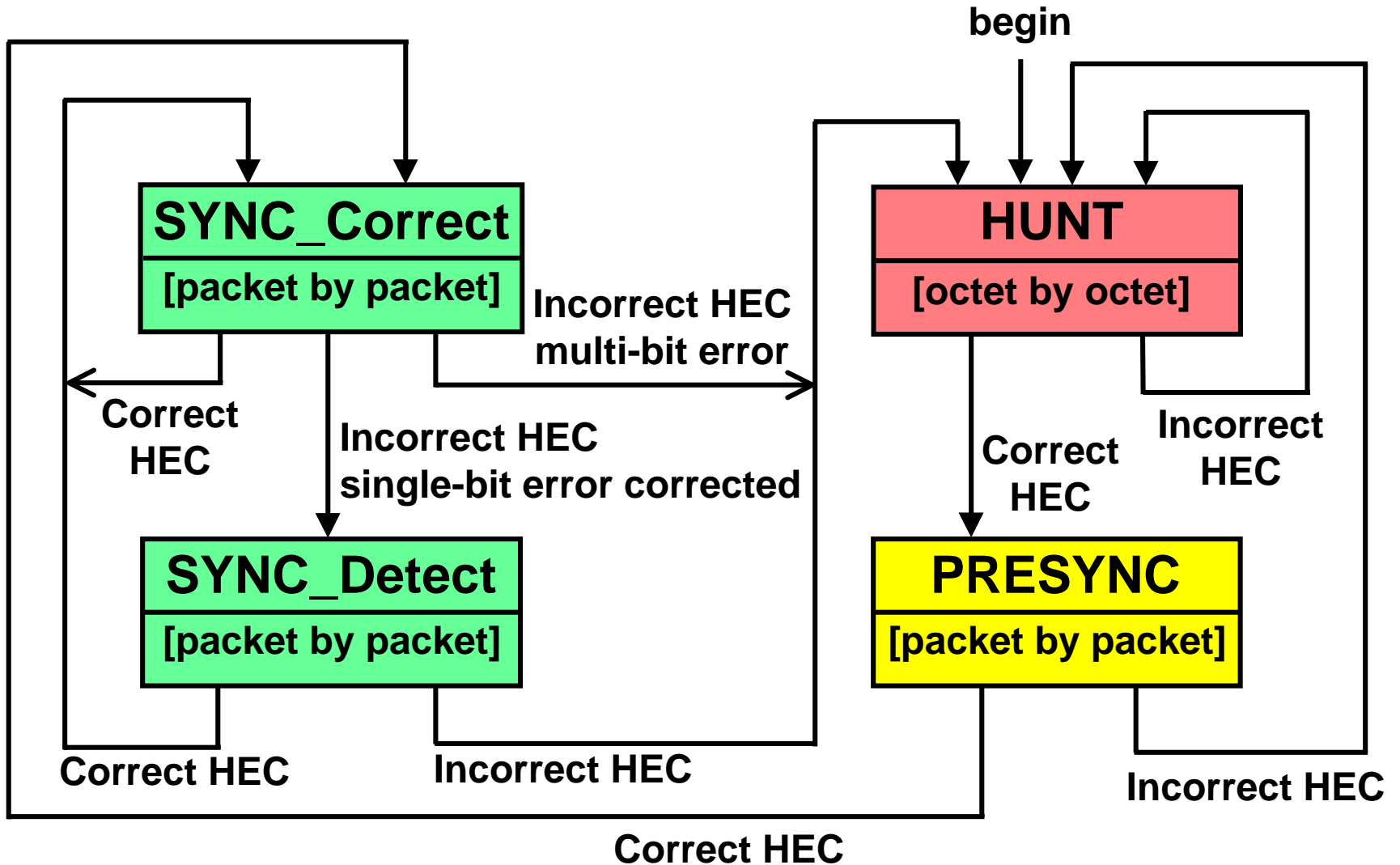
Packet Delineation: PRESYNC state



Packet Delineation: SYNC states



Complete State Diagram

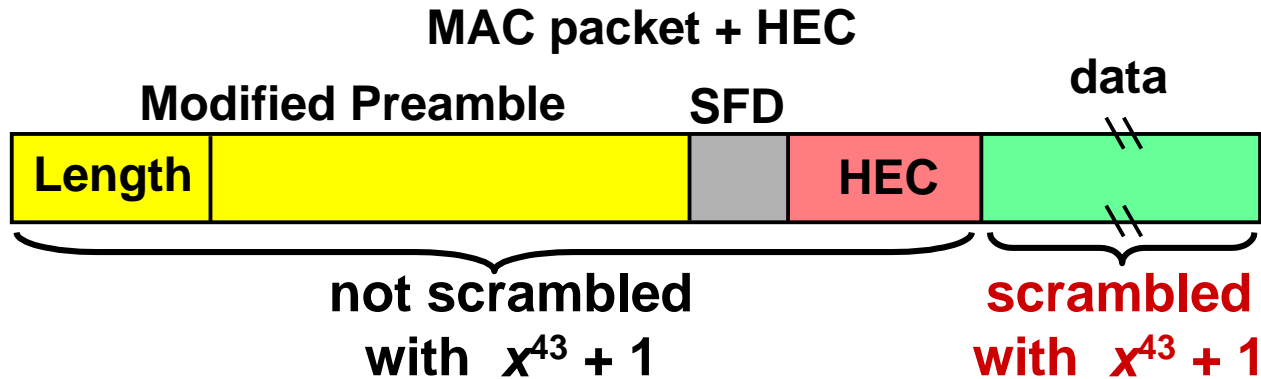


Packet Delineation Performance

- **Reference**
 - “10GE WAN PHY Delineation Performance”
 - http://grouper.ieee.org/groups/802/3/10G_study/public/email_attach/delineation_perf.doc
- **Probability of frame loss = 9.56×10^{-21}**
 - Frame loss = state machine moves from SYNC_Correct or SYNC_Detect to the HUNT state
 - Assumption: BER = 10^{-12}
- **Mean Time to Frame Loss = 0.7 million years**
 - Average frame length of 500 bytes

$x^{43}+1$ Self-Synchronous Scrambler

- PCS scrambles all the octets between the HEC field (excluded) and the end of the MAC packet
 - i.e., only <data> is scrambled
 - Idle PHY packets are not scrambled



Purpose of $x^{43}+1$ Scrambler

- **Protects DC Balance**

- Malicious users cannot transmit “killer packets” to disrupt DC balance (which is provided by the $x^7 + x^6 + 1$ scrambler at the PMA sublayer)
- Killer packets can be generated only if user knows the state of the scrambler. The odds of guessing it correctly are 1 in 2^{43}

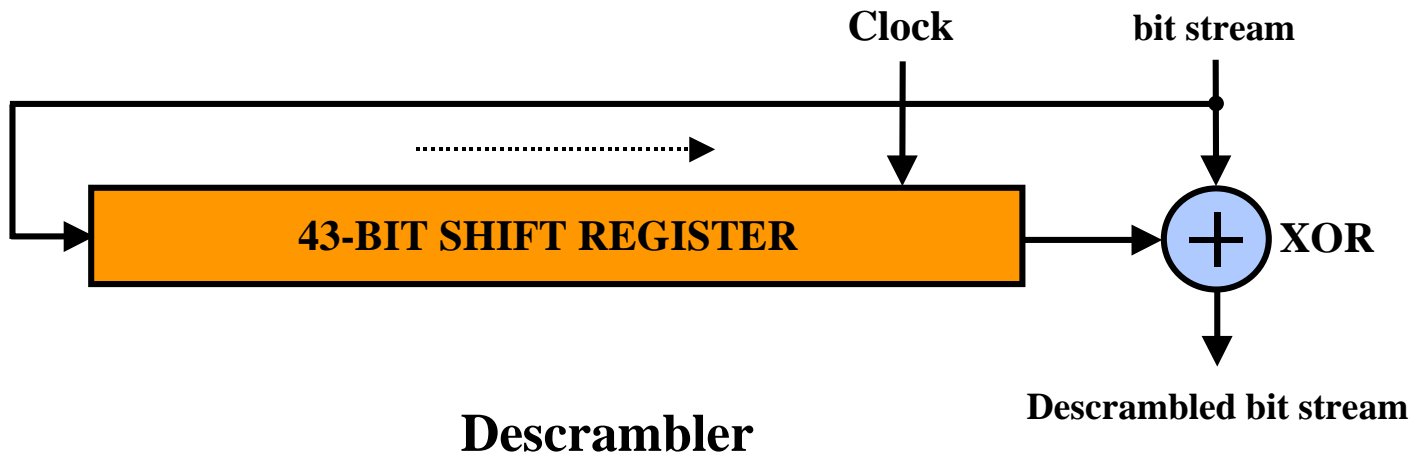
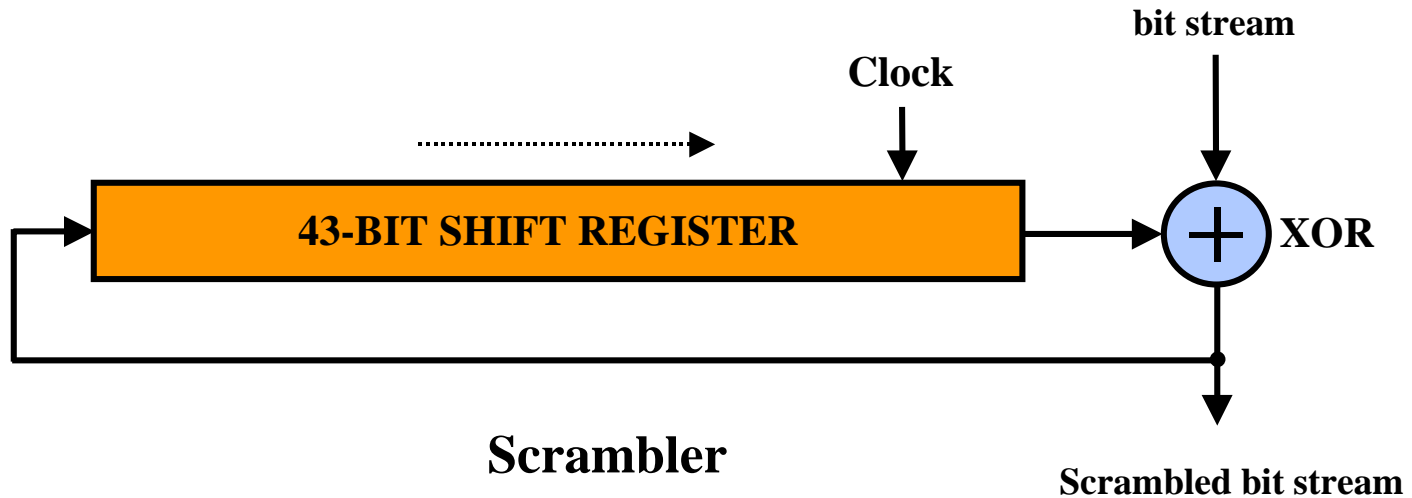
- **Protects packet delineation**

- packet delineation is done on the scrambled octet stream
- Malicious users cannot emulate HEC fields

- **Note:**

- Packet delineation is not affected by the error duplication effect of the $x^{43} + 1$ scrambler

$x^{43}+1$ Scrambler/Descrambler



$x^{43}+1$ Scrambler/Descrambler State

- **Content of the 43-bit shift register is retained when the scrambler/descrambler is disabled**
- **Initial state after power on or main reset**
 - of the scrambler should be randomly selected to improve security
 - of the descrambler is irrelevant
- **Scrambling is enabled**
 - Only from first octet of DA to last octet of FCS
- **Descrambling is enabled**
 - Only in the SYNC_Correct and SYNC_Detect states
 - Only from first octet of DA to last octet of FCS

Bit Order of $x^{43}+1$ Scrambling

- **Least significant bit (LSB) first**
 - To agree with FCS calculation
 - If scrambling and FCS calculation were reversed with respect to each other, the short burst error detection capabilities of the FCS would be degraded

