10GE WAN PHY Overview (or UniPHY WIS)

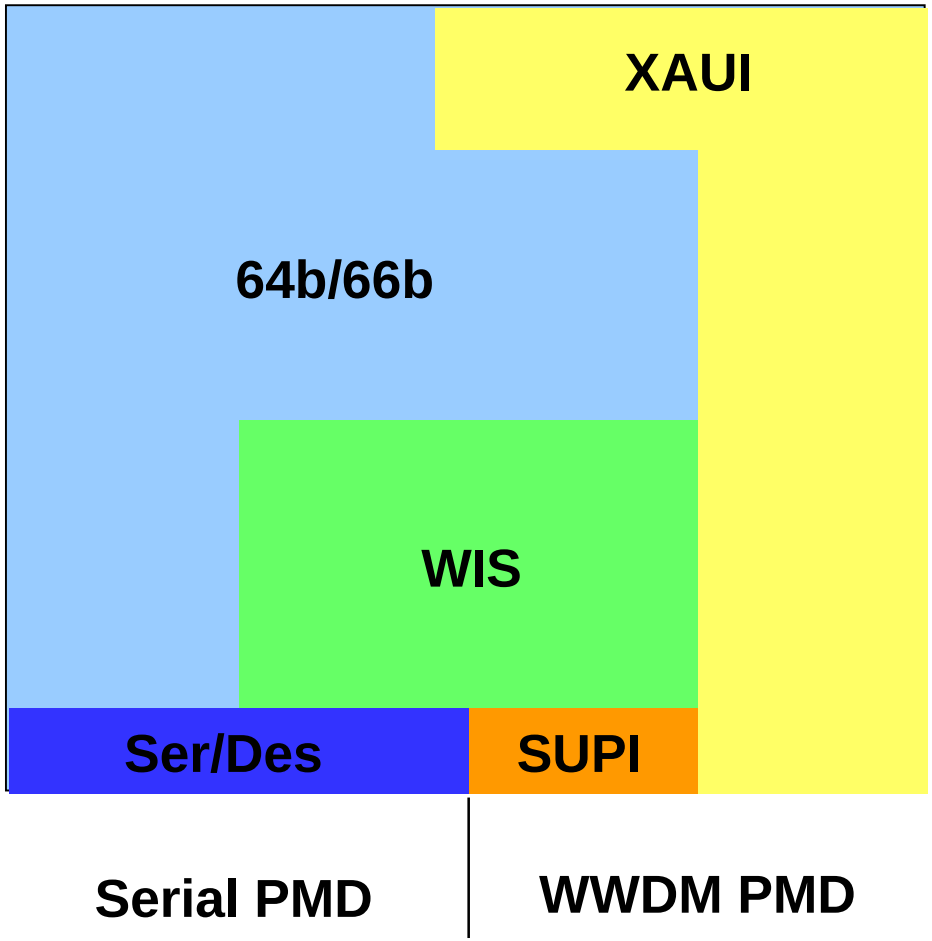
IEEE 802.3ae Meeting, Ottawa
May 23-25, 2000

Norival Figueira, Paul Bottorff, David Martin, Tim Armstrong, Bijan Raahemi: Nortel Networks - Howard Frazier: Cisco Systems - Enrique Hernandez: Lucent (Bell Labs) - Nevin Jones: Lucent Microelectronics - Steve Haddock: Extreme Networks - Pankaj Kumar, Bradley Booth: Intel - Bjørn Liencres: Juniper Networks - Tom Palkert: AMCC - Iain Verigin, Stuart Robinson, Tom Alexander : PMC Sierra - Nader Vijeh: Lantern Communications - Frederick Weniger: Vitesse - Shimon Muller: Sun Microsystems

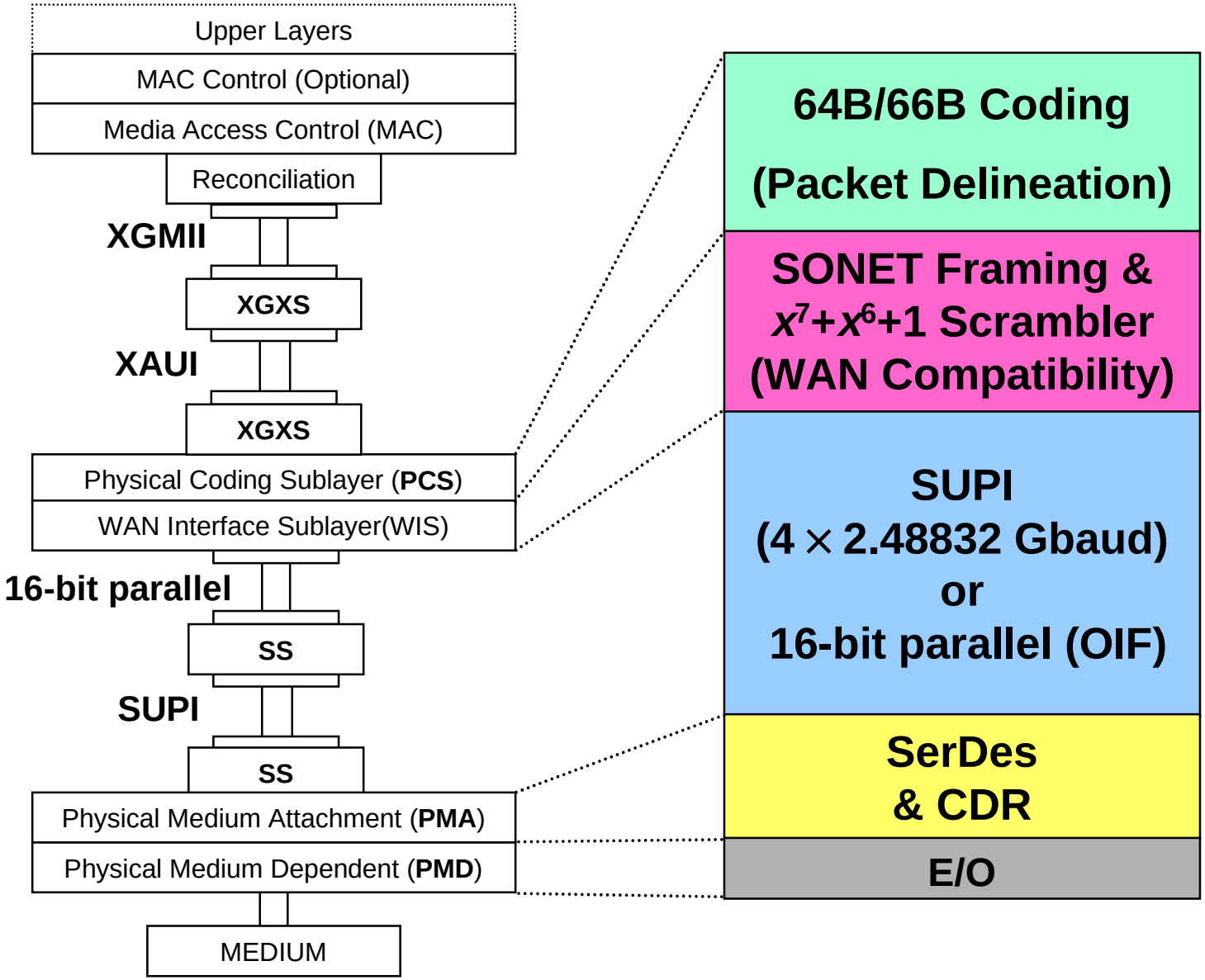
Agenda

- **Supported PMDs**
- **PMD interfaces**
  - SUPI
  - 16-bit Parallel
- **WIS (WAN Interface Sublayer)**
  - SONET frame and overheads
  - WIS framing function
  - WIS frame synchronization process
  - $x^7 + x^6 + 1$ frame-synchronous scrambler
- **PCS**
  - PCS (Packet Delineation)

WAN-PHY and LAN-PHY Components

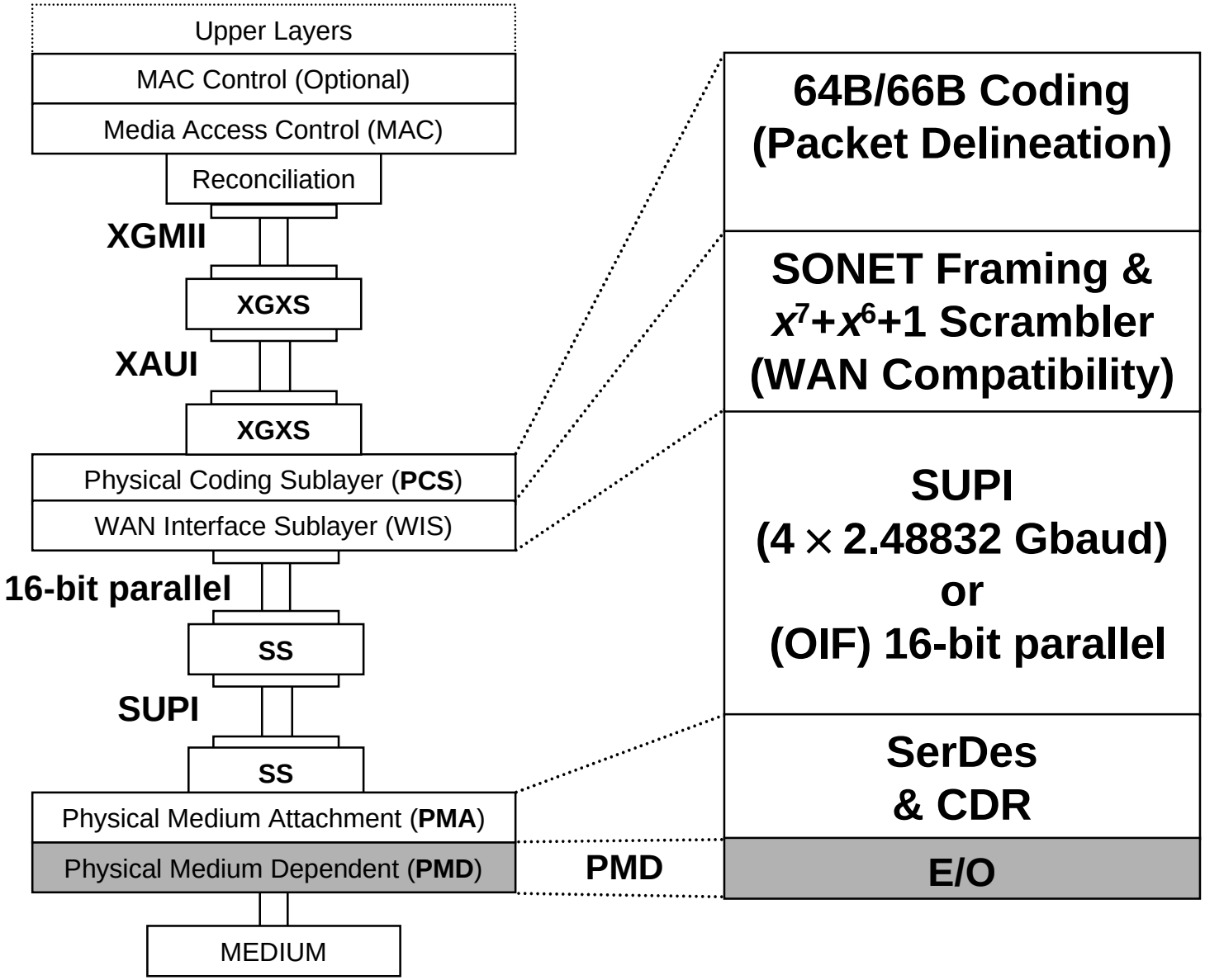


WAN-PHY Layer Model



SS = SUPI Sublayer

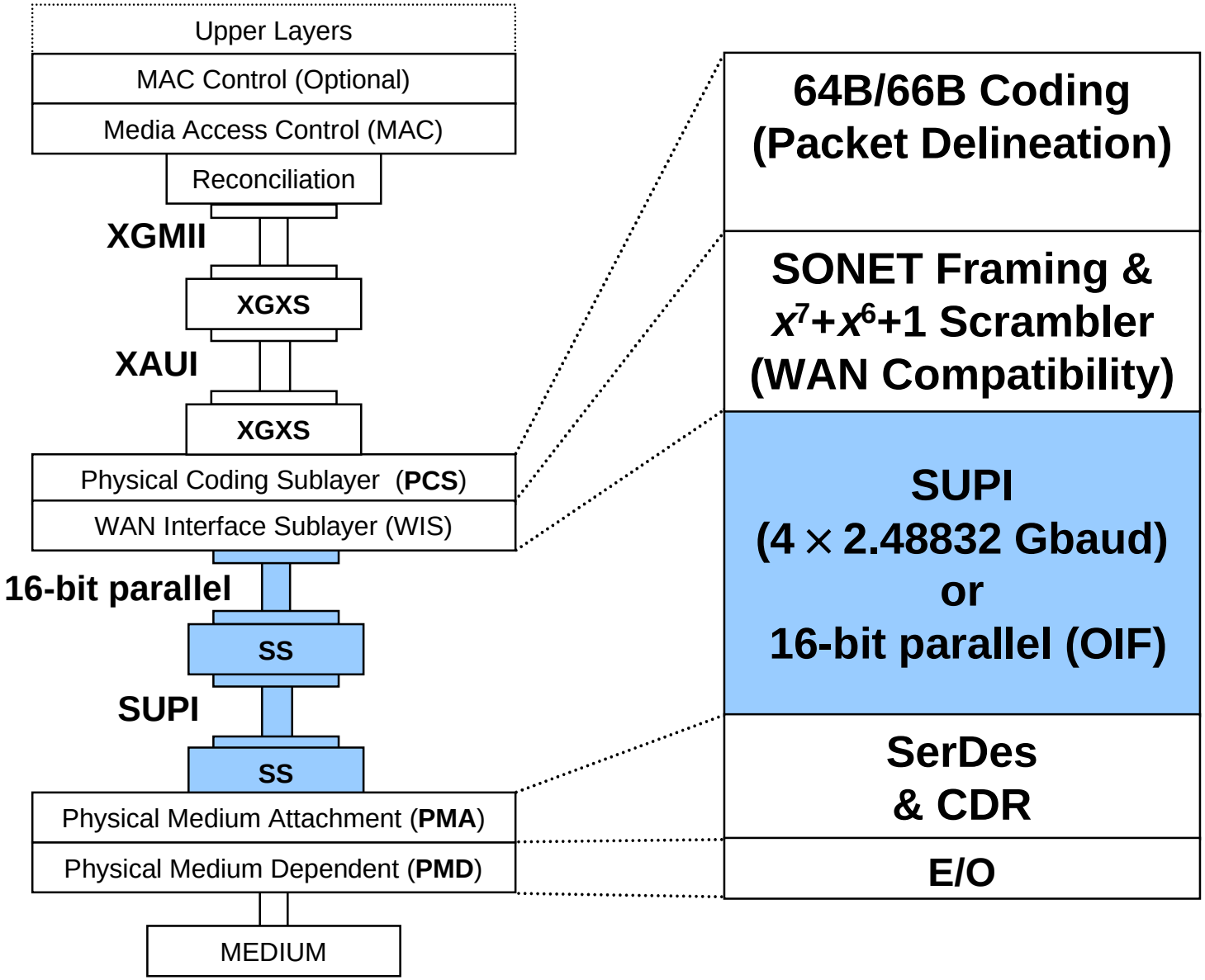
PMD


Upper Layers
MAC Control (Optional)
Media Access Control (MAC)
Reconciliation
XGMII
XGXS
XAUI
XGXS
Physical Coding Sublayer (PCS)
WAN Interface Sublayer (WIS)
16-bit parallel
SS
SUPI
SS
Physical Medium Attachment (PMA)
Physical Medium Dependent (PMD)
MEDIUM
PMD
64B/66B Coding
(Packet Delineation)
SONET Framing &
x7+x6+1 Scrambler
(WAN Compatibility)
SUPI
(4 × 2.48832 Gbaud)
or
(OIF) 16-bit parallel
SerDes
& CDR
E/O

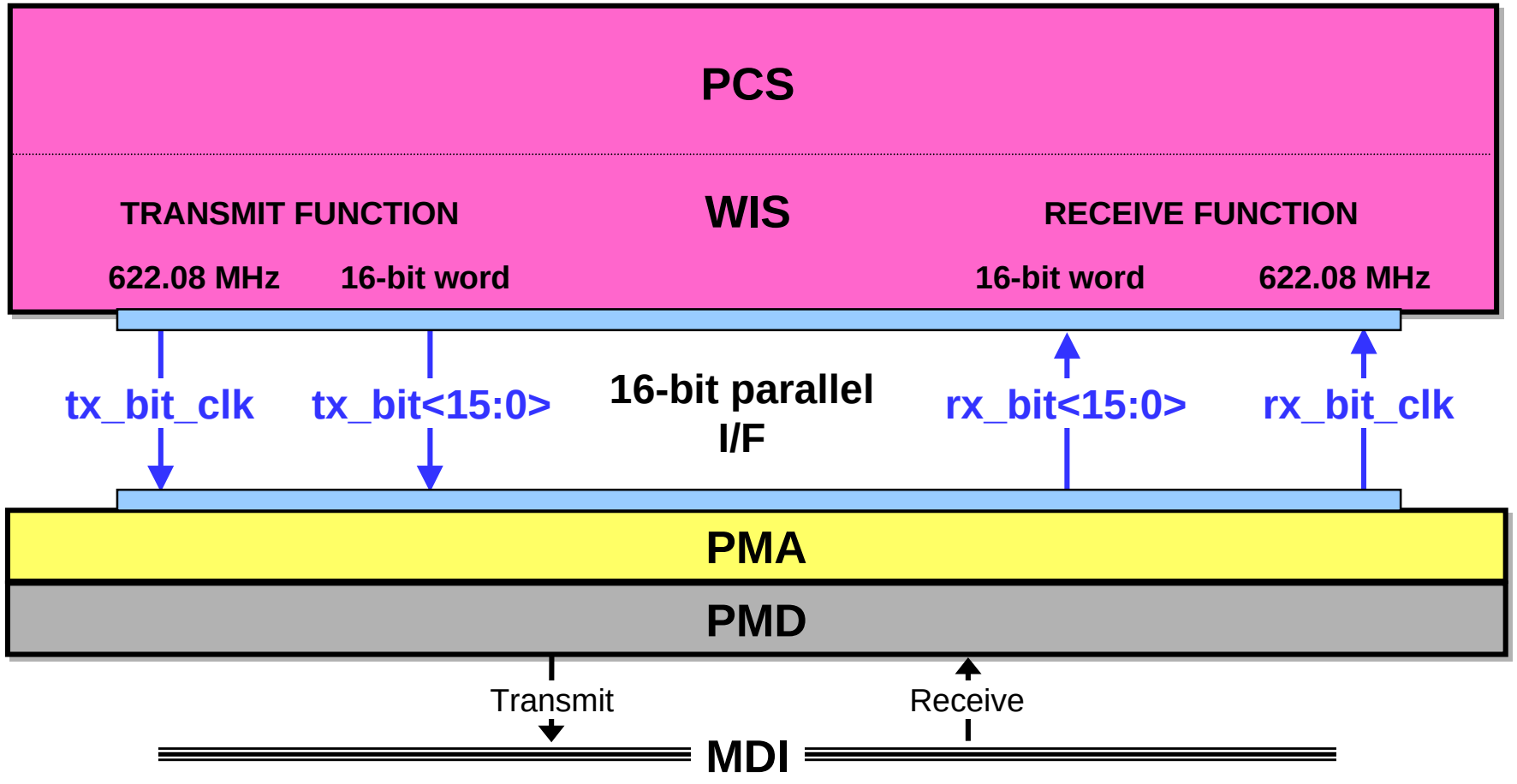
Need all the PMDs on the Short List

- **Main application of WAN PHY is short distance…**
  - High volume tera-POP applications
  - Point to Point SONET Interconnect and DWDM case
- **However, longer distances are also important**
  - Dark fiber
  - MANs
- **Conclusion:**
  - WWDM PMDs must also support the WAN PHY

PMD Interfaces



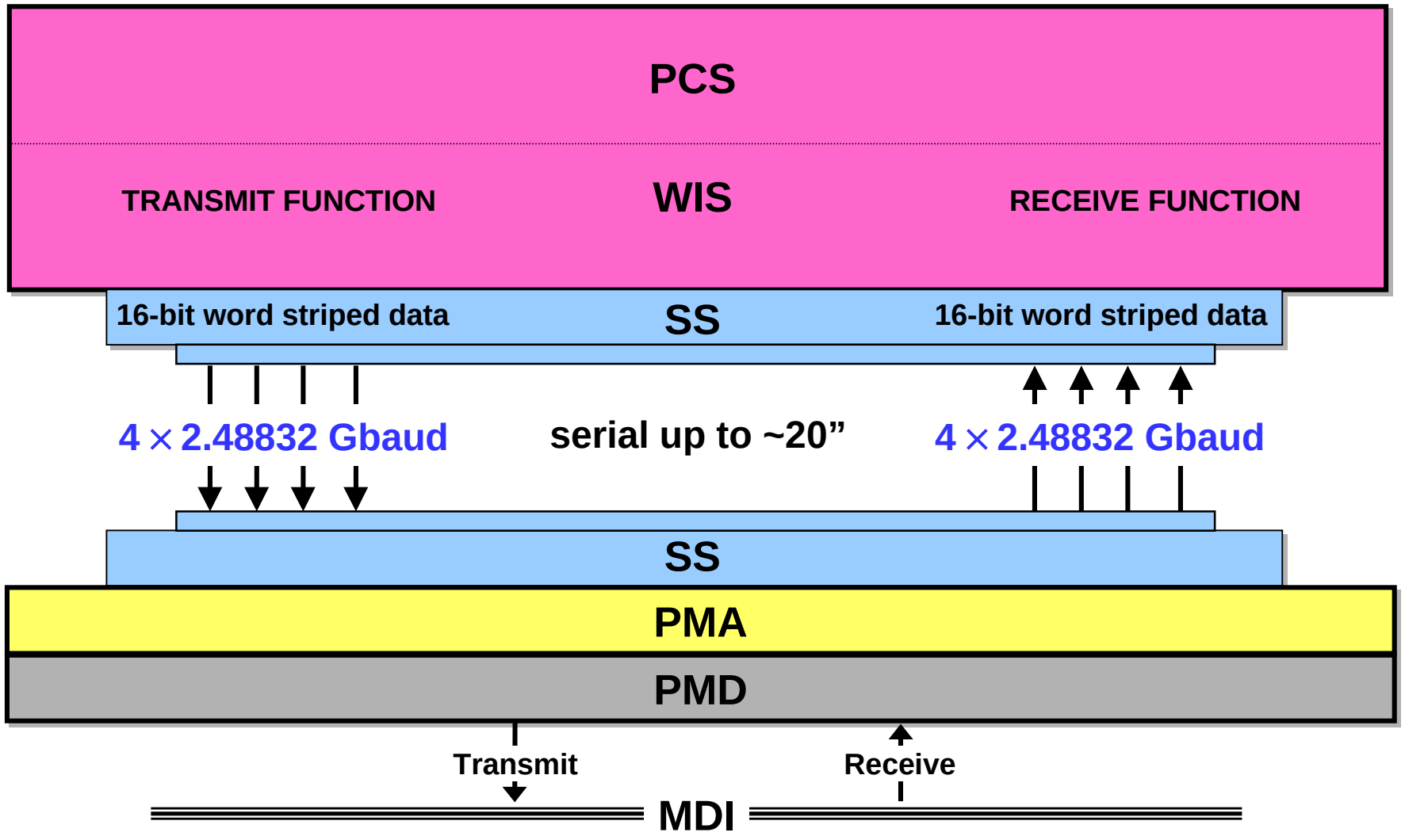
16-Bit Parallel PMD Interface



16-bit Parallel PMD Interface (cont.)

- **tx_bit<15:0>**
  - 16-bit vector representing two octets received from the WIS
  - transitions synchronously with tx_bit_clk
- **tx_bit_clk**
  - 622.08 MHz clock generated by the WIS
- **rx_bit<15:0>**
  - Most recently received 16 bits (MSB first) from the MDI. It is a continuous and unaligned sequence of octets
  - transitions synchronously with rx_bit_clk
- **rx_bit_clk**
  - 622.08 MHz clock generated by the PMD
- **all LVDS**

SUPI (WWDM PMD Interface)



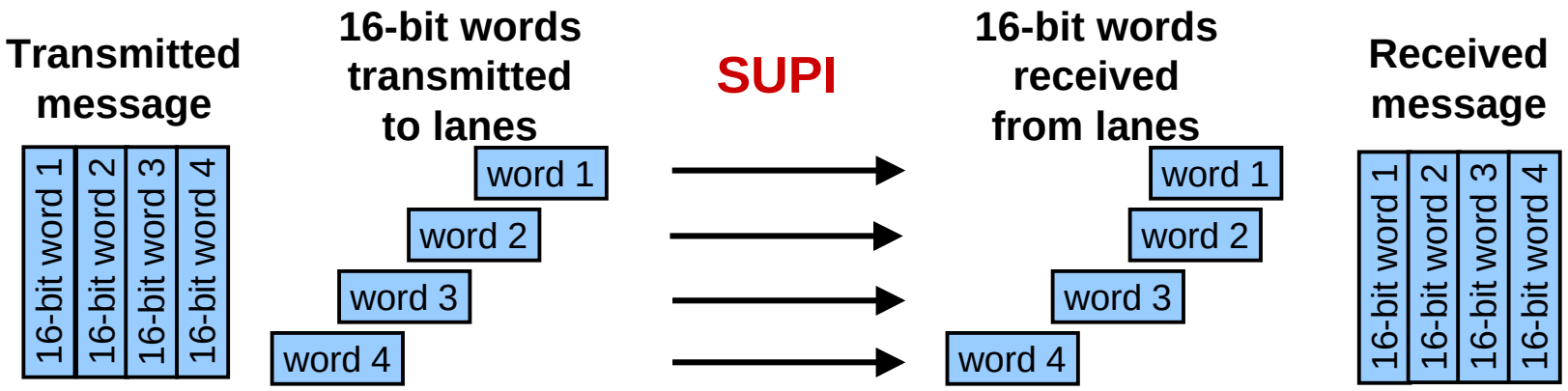
SS = SUPI Sublayer

SUPI

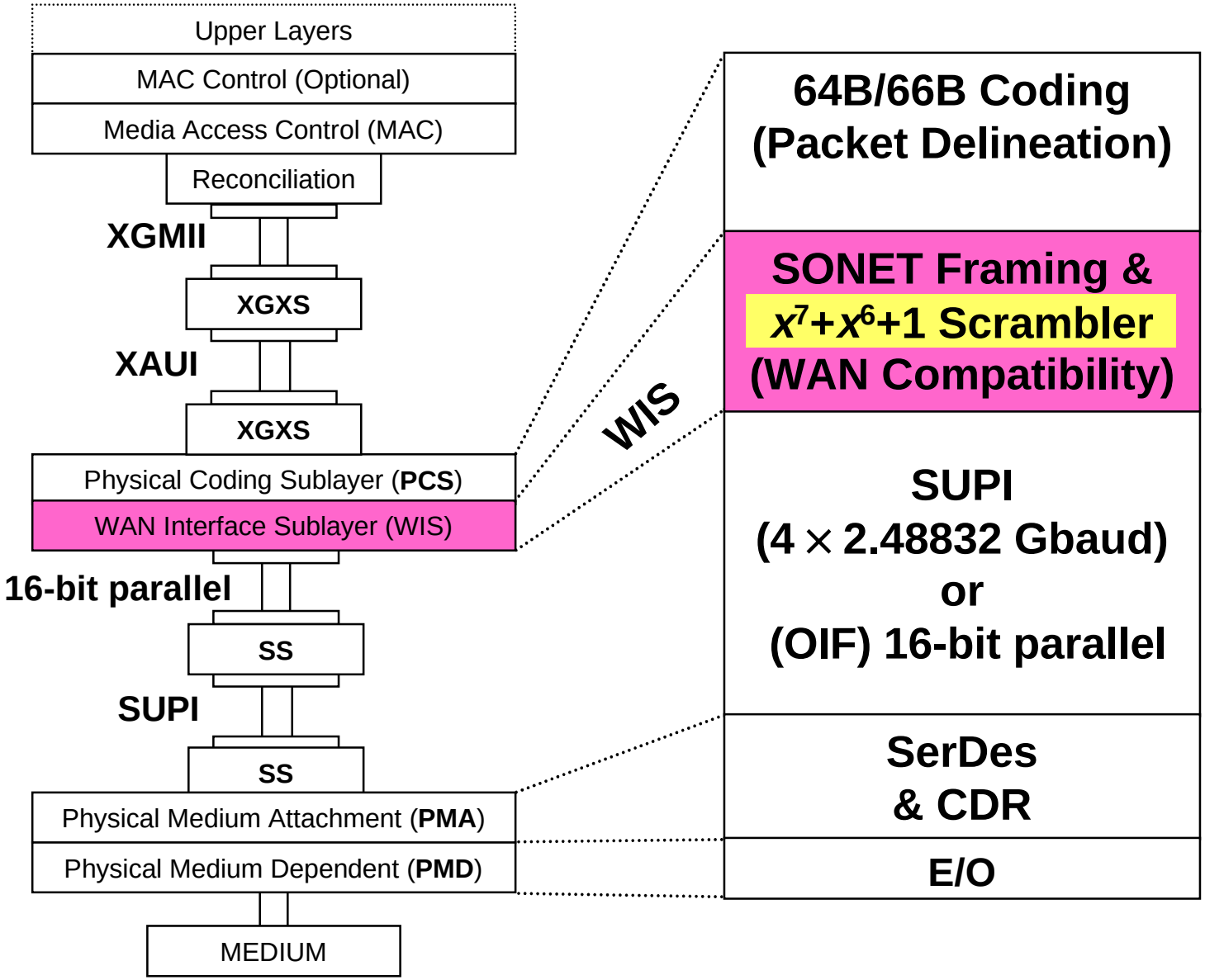
- **Same physical interface as XAUI...**
  - 4 serial lanes
  - Up to ~20”
- **...But uses different encoding and clock rate**
  - 2.48832 Gbaud per lane
  - Scrambled code
- **Used for WWDM and 4× parallel PMDs**
  - Must work with scrambled code at 2.48832 Gbaud (SUPI) and block code at 3.125 Gbaud (XUAI)
  - Can use a recovered clock to reset jitter

SUPI (cont.)

- **16-bit word striped data transmitted on each lane**
- **Each lane has 1/4 of the (SONET) A1/ A2 framing bytes for lane deskew and synchronization**
  - Word synchronization from A1/A2 transition
  - For fixed lane assignment, allows for large skew

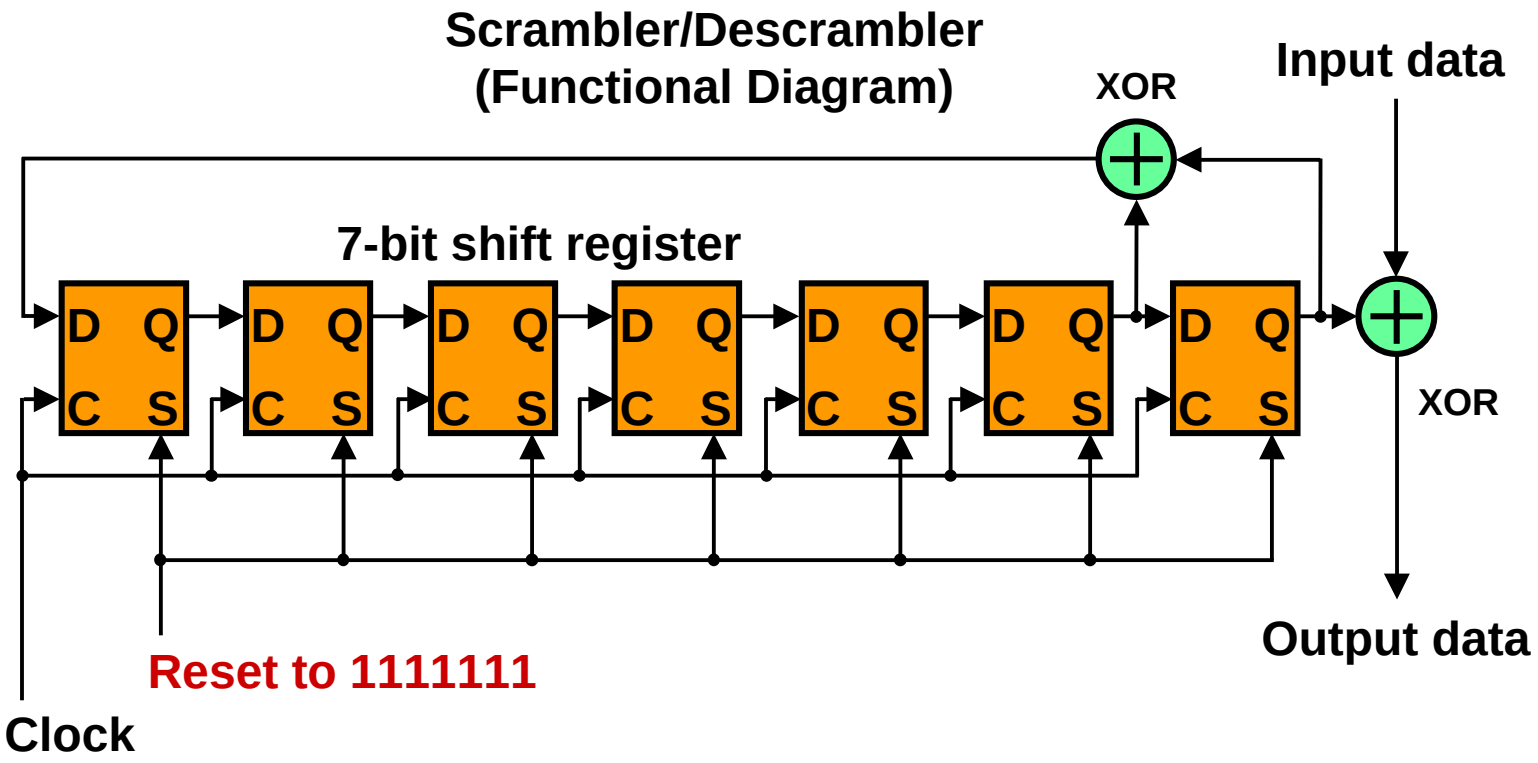


x^7+x^6+1 Scrambler

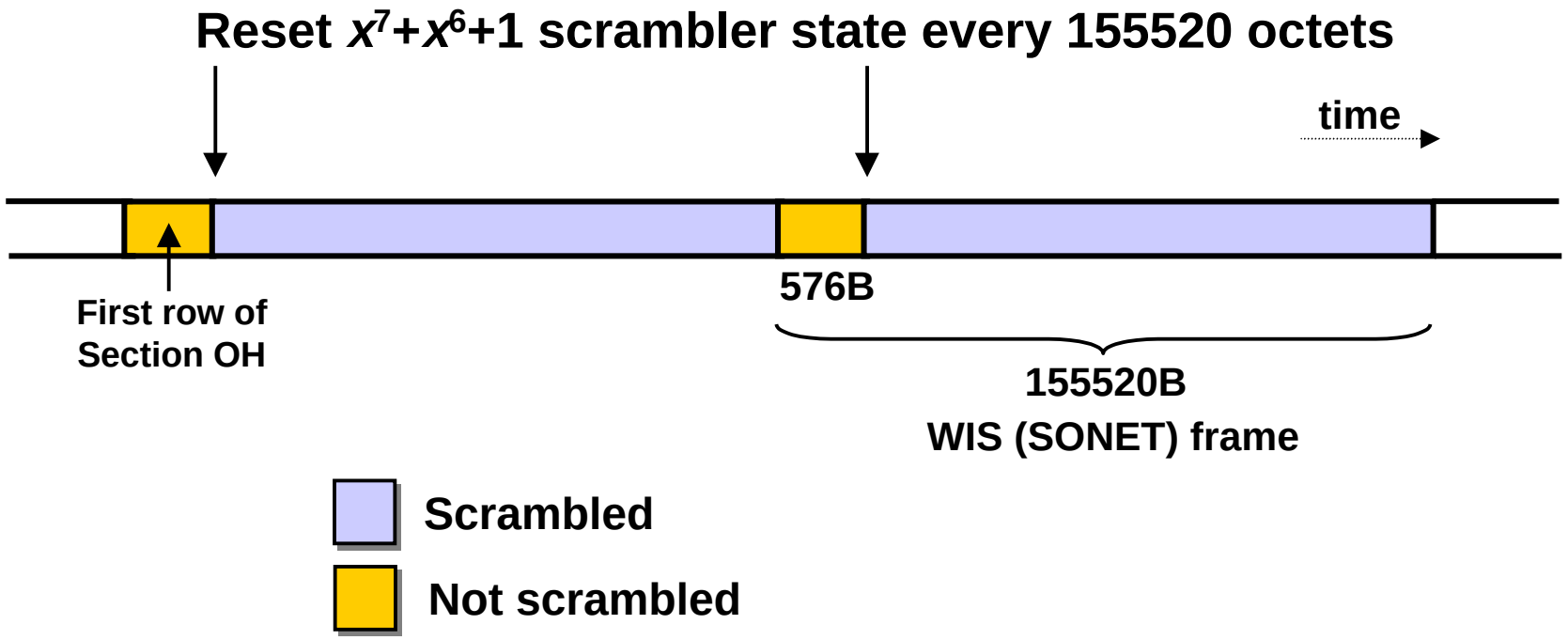


x^7+x^6+1 Scrambler (cont.)

- **Provides high randomization**
  - Assures adequate number of transitions for line rate clock recovery at the receiver

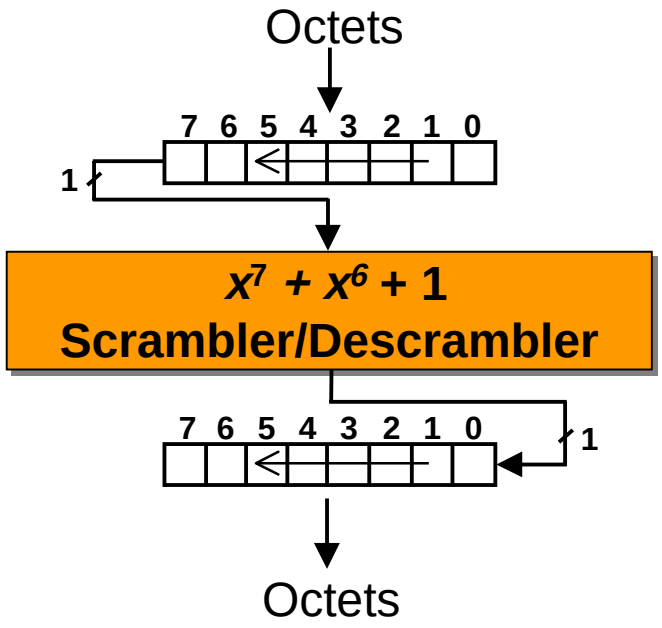


State is Periodically Resynchronized



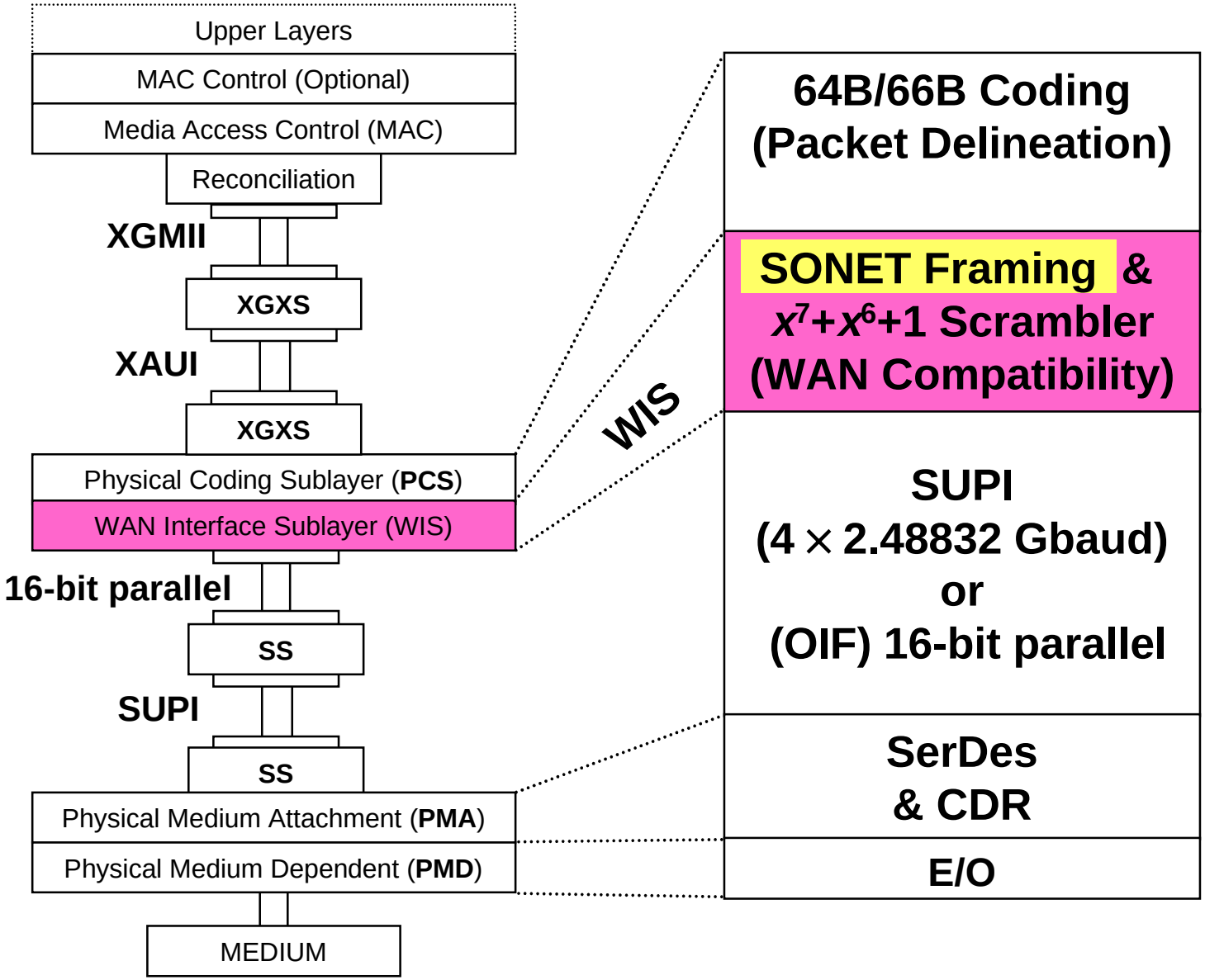
Bit Order of Scrambling/Descrambling

- **Most significant bit (MSB) first**



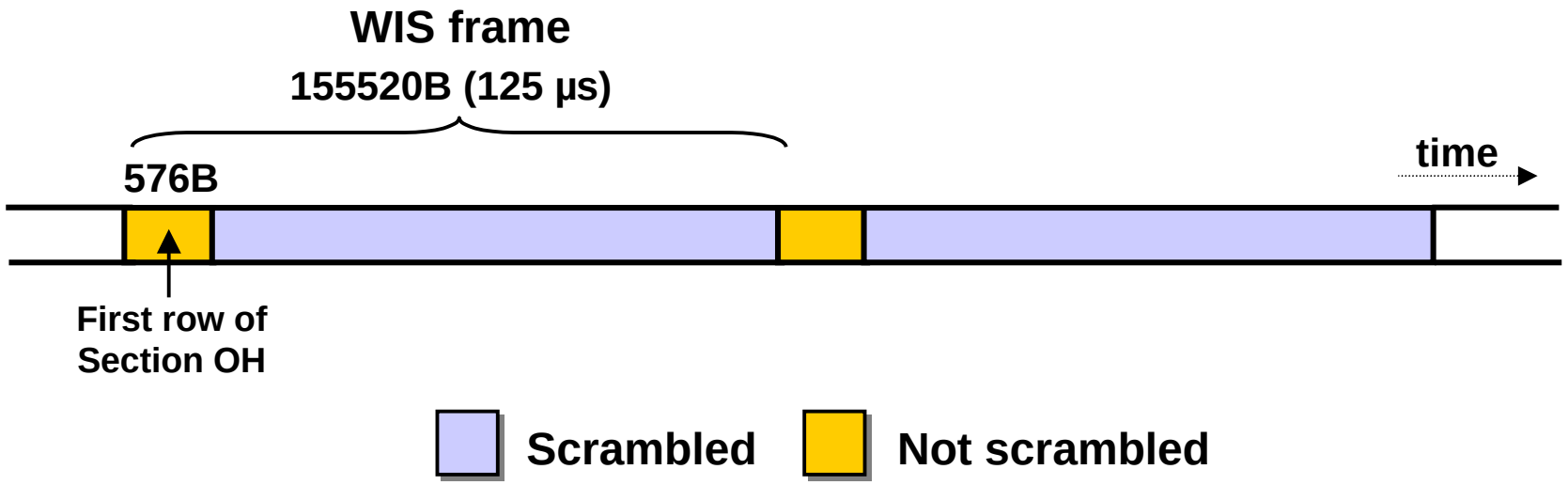
(Functional diagram)

SONET Framing

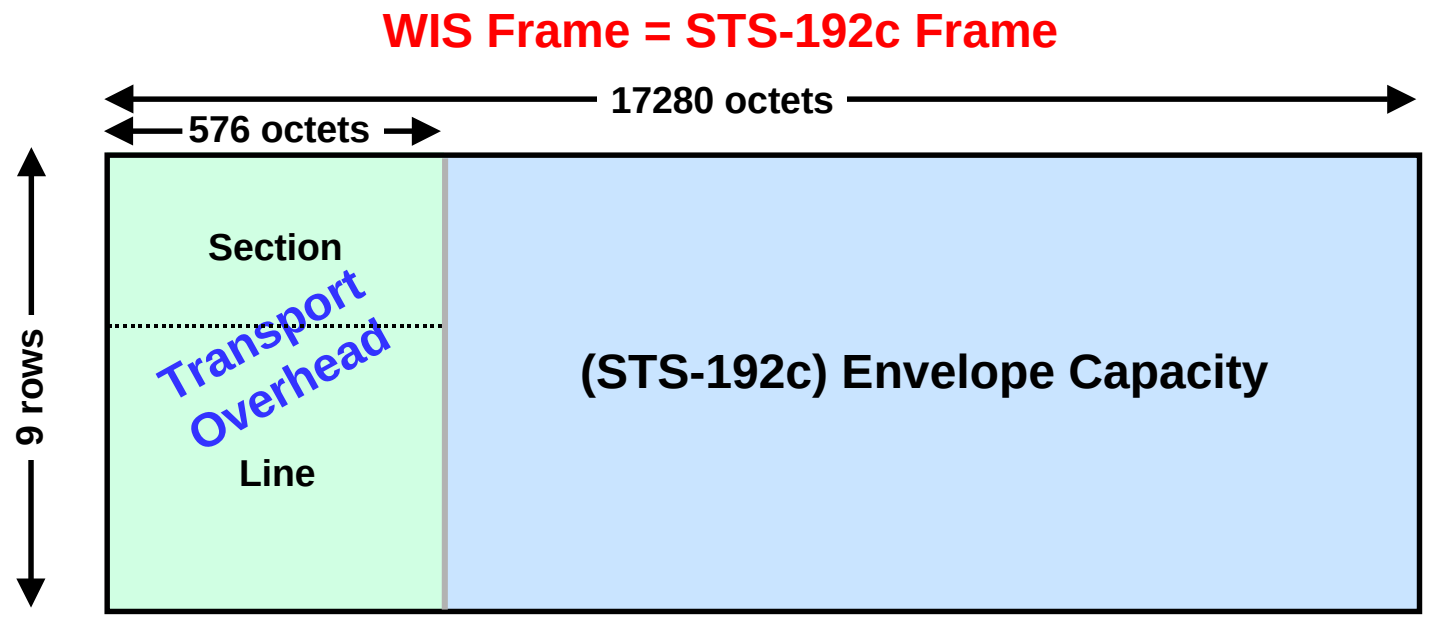


WIS Frame

- **SONET frame with minimum overhead support**
  - **Overheads are out of band management used to control SONET networks**
  - **While the WIS frame is compatible with SONET, it does not provide full SONET management**
- **Sequence of 155520 octets (125 µs)**



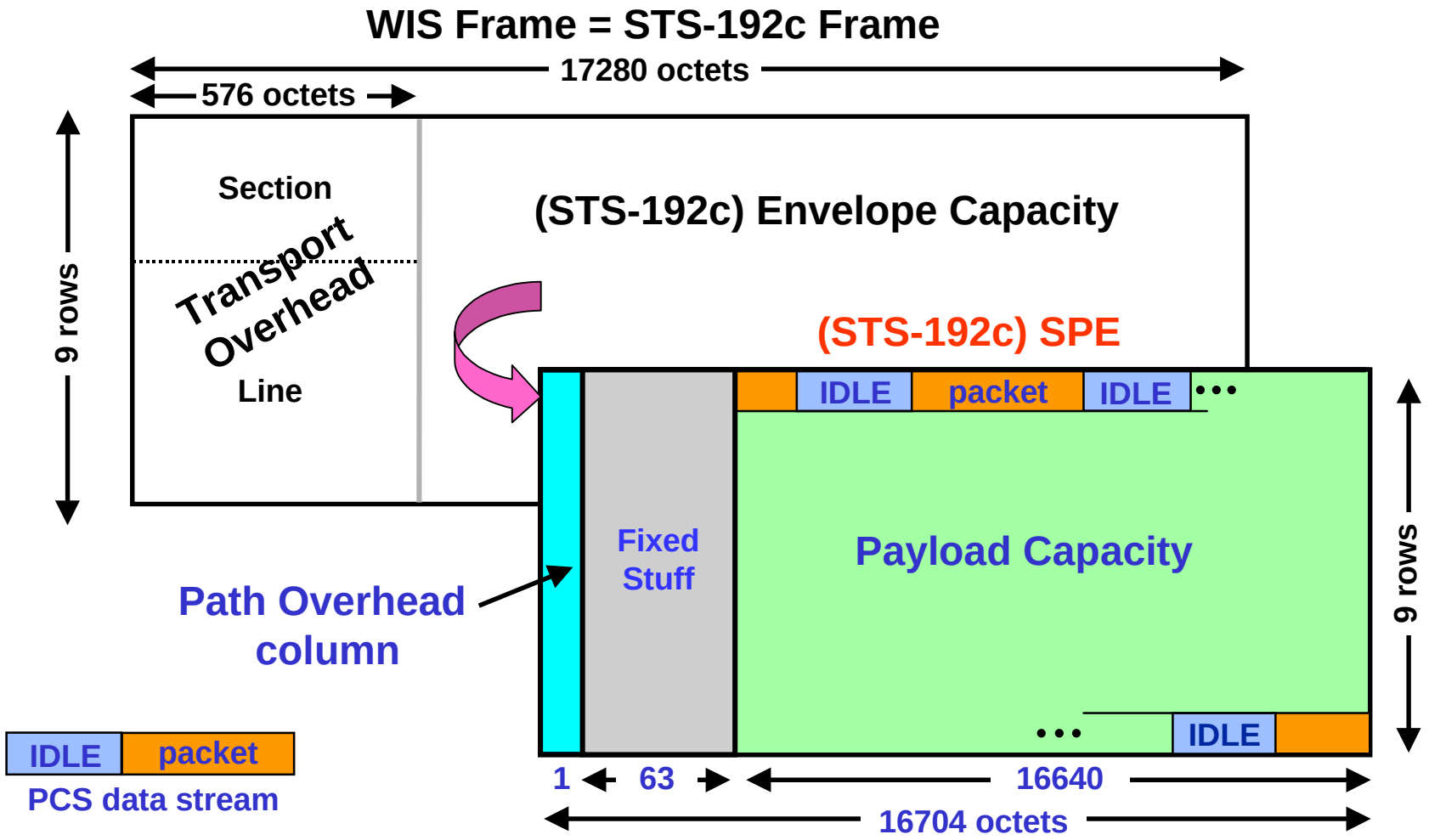
WIS Frame: Viewed as 9×17280 Octets



STS-192c = Synchronous Transport Signal – level 192, c = concatenated.

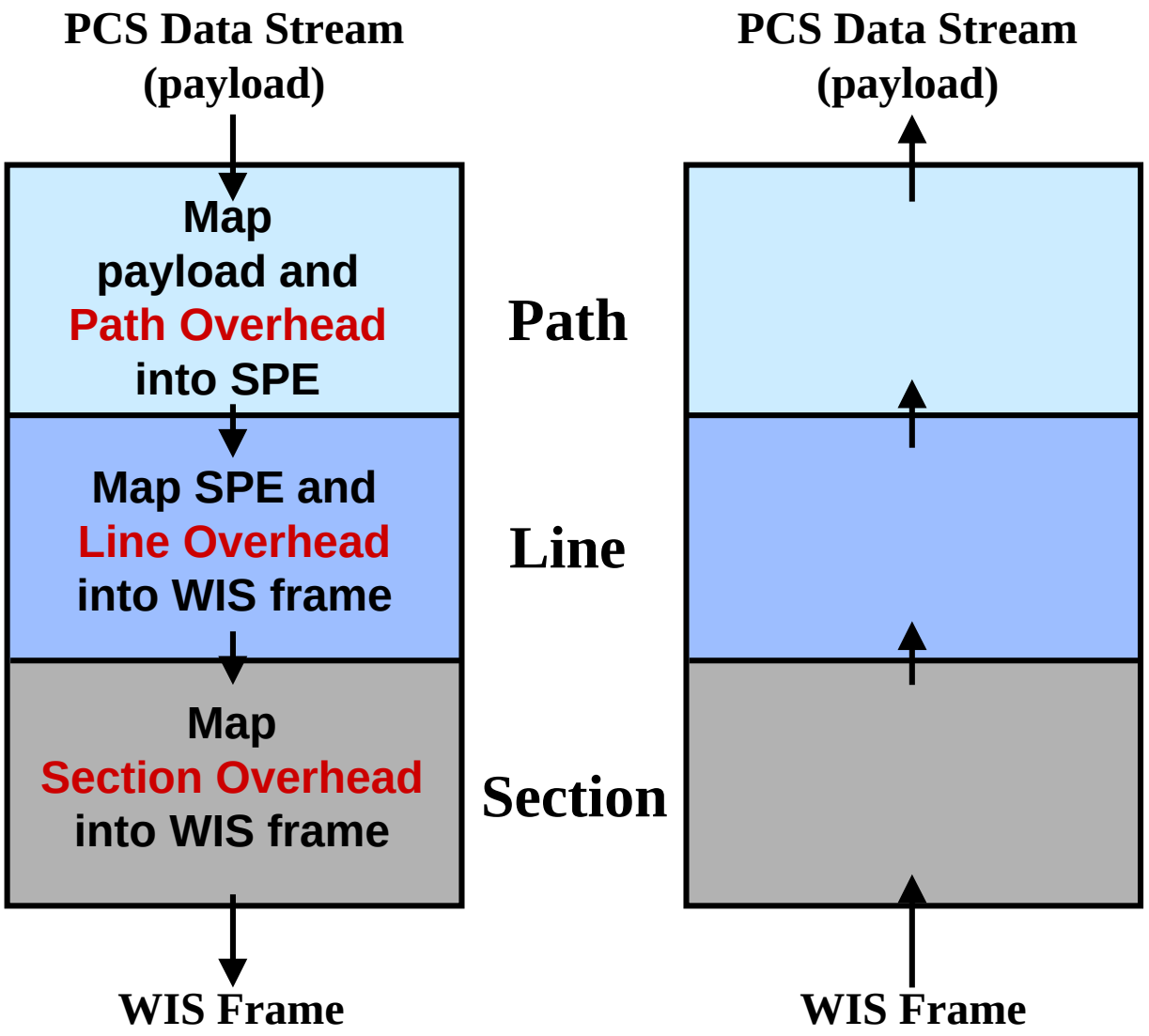
Transmission order: top to bottom, row-by-row, left to right.

Payload Capacity (9.58464 Gb/s)

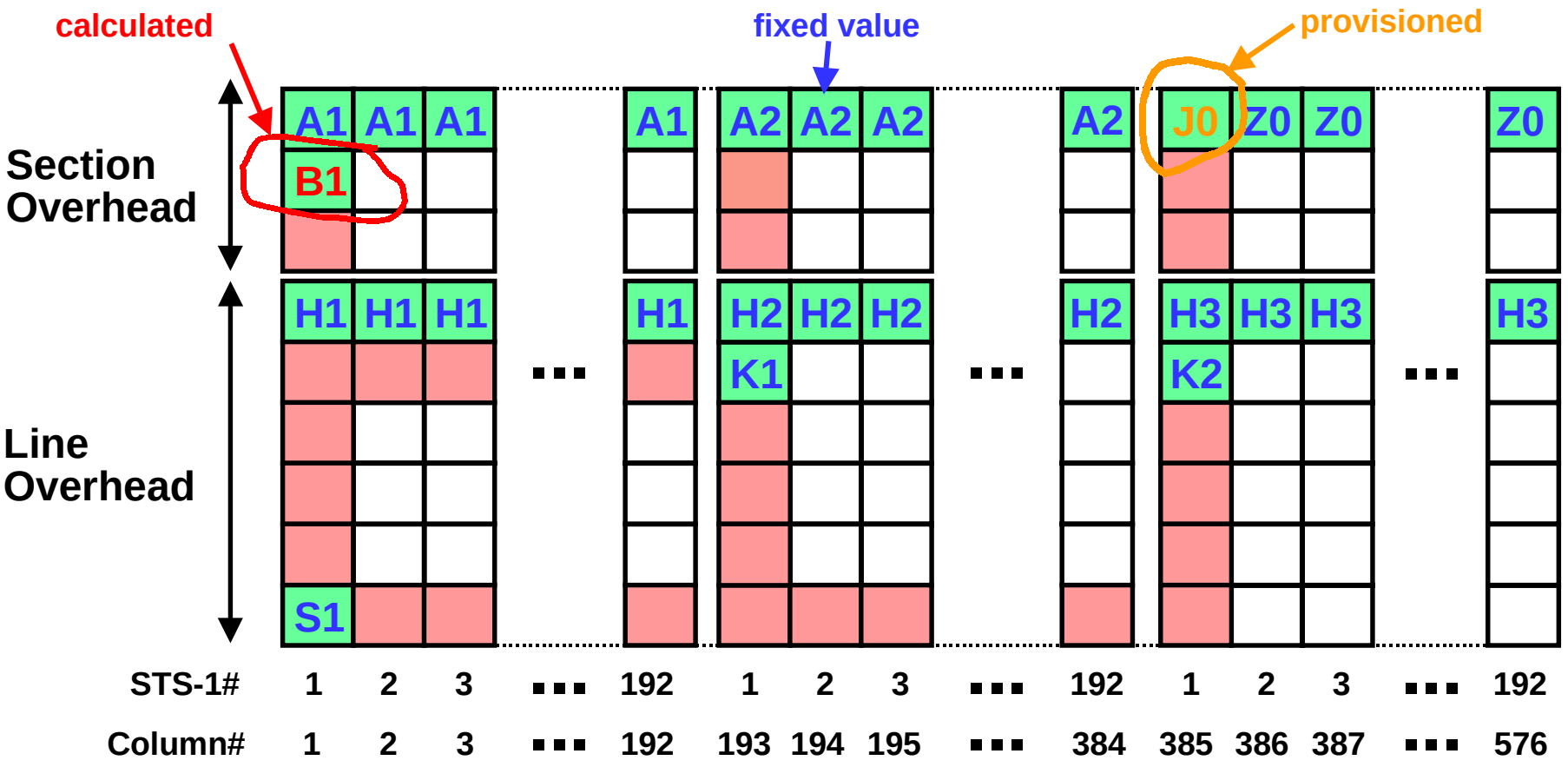


STS-192c = Synchronous Transport Signal – level 192, c = concatenated
SPE = Synchronous Payload Envelope

WIS Overhead Layers



Transport Overhead



= Undefined overhead octets (set to zero)

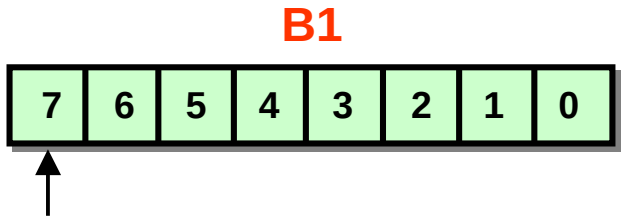
= Defined overhead octets (B2, E1-2, F1, D1-12, M1, Z1-2), unused by 10GE WAN PHY (set to zero)

Section Overheads

- **A1 and A2 (“Framing octets”)**
  - Fixed value: A1 = 11110110, A2 = 00101000
  - A1/A2 transition is used for WIS frame synchronization
- **J0 (“Section Trace”)**
  - Allows a receiver to verify its continued connection to the intended transmitter
  - Provisioned Value
    - when no value is provisioned, J0 shall be set to 00000001)
- **Z0 (‘Section Growth”)**
  - Fixed value: 11001100

Section Overheads (cont.)

- **B1 (“Section BIP-8”)**
  - — Used as a Section error monitoring function
  - — Calculated value:
    - – BIP-8 code (using even parity) over all the bits of the last transmitted WIS frame <u>after</u> scrambling

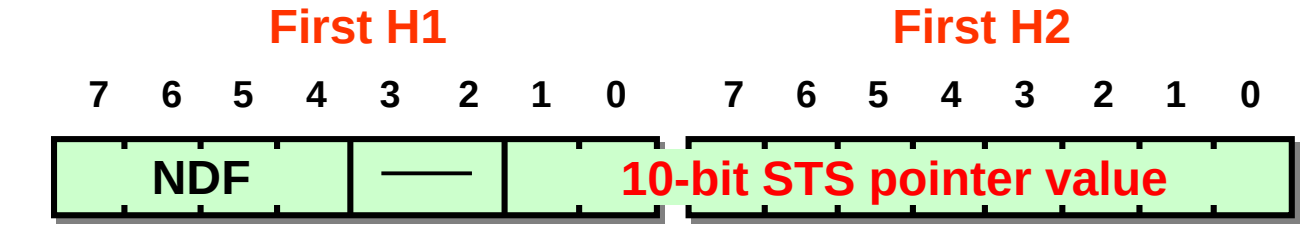


BIP-8 (Bit-Interleaved Parity-8) with even parity: The i^{th} bit of the code provides even parity over the i^{th} bit of all the covered octets.

BIP-8of the bit sequence 11110000 00001111 is 11111111.

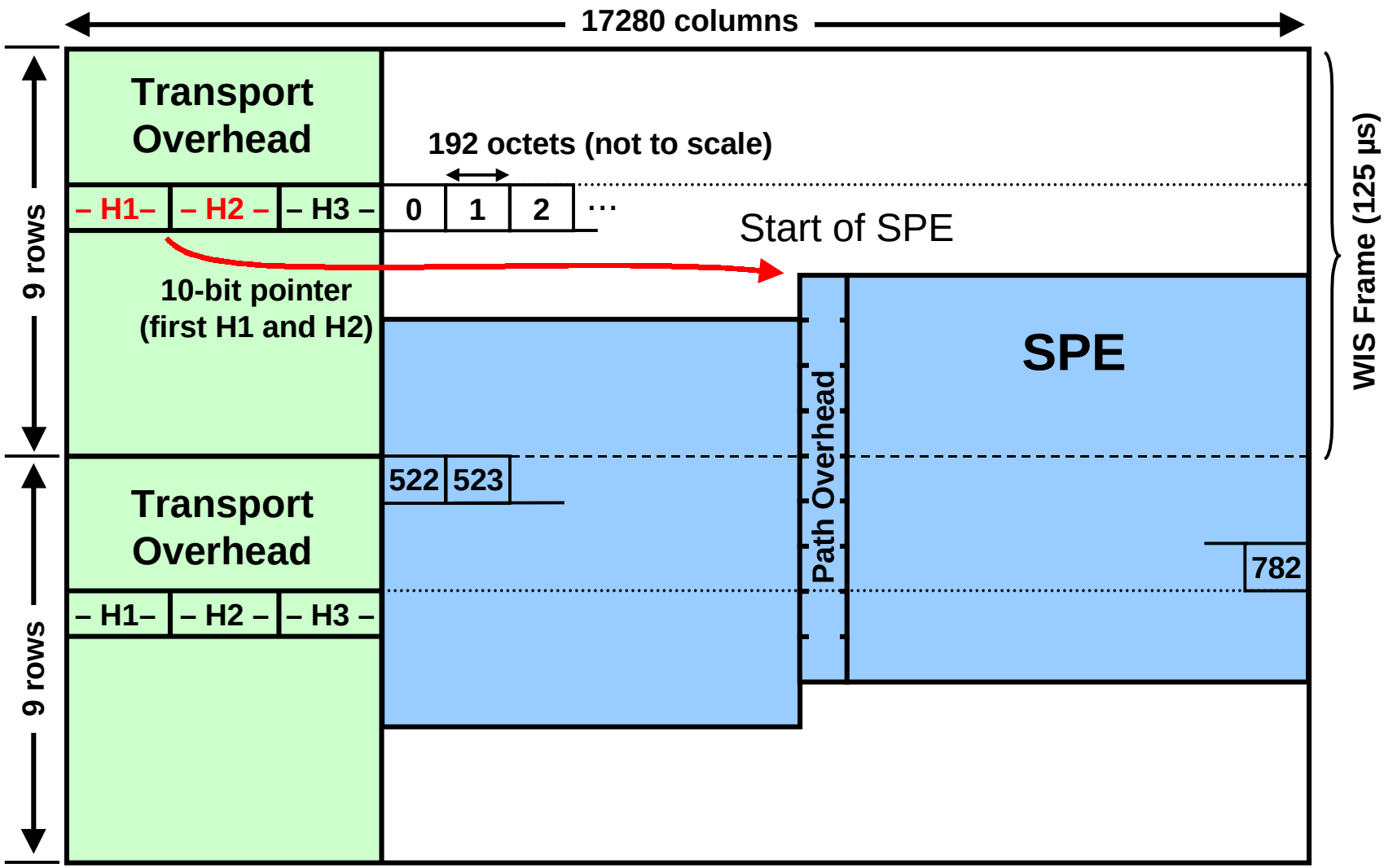
Line Overheads

- **First H1 and H2 (“Payload Pointer”)**
  - 16-bit word containing 10-bit pointer in the range of 0 to 782
  - Transmits fixed values: H1 = 01100010 and H2 = 00001010 (i.e., pointer = 522)
  - Receiver 10GE WAN PHY shall be able to process arbitrary pointer values (which may be changed by a transport network)
- **Second to last H1 and H2**
  - Fixed Values: H1 = 10010011 and H2 = 11111111



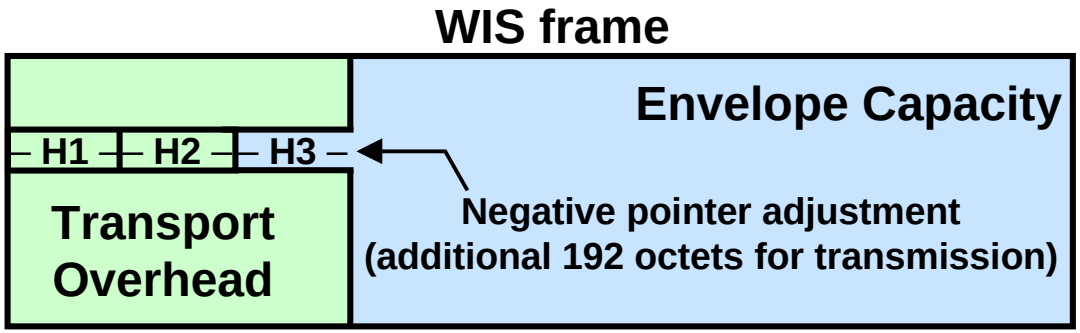
NDF (new data flag) field

H1/H2 Pointer and SPE Position



Line Overheads (cont.)

- **H3 (“Pointer Action Bytes”)**
  - Allows an LTE to have slightly different clocks at the receiver and transmitter paths
  - Carries 192 extra SPE (payload) octets in the event of a “negative pointer adjustment,” which may be required when the receiver clock is faster than the transmitter clock
  - Set to zero when not used

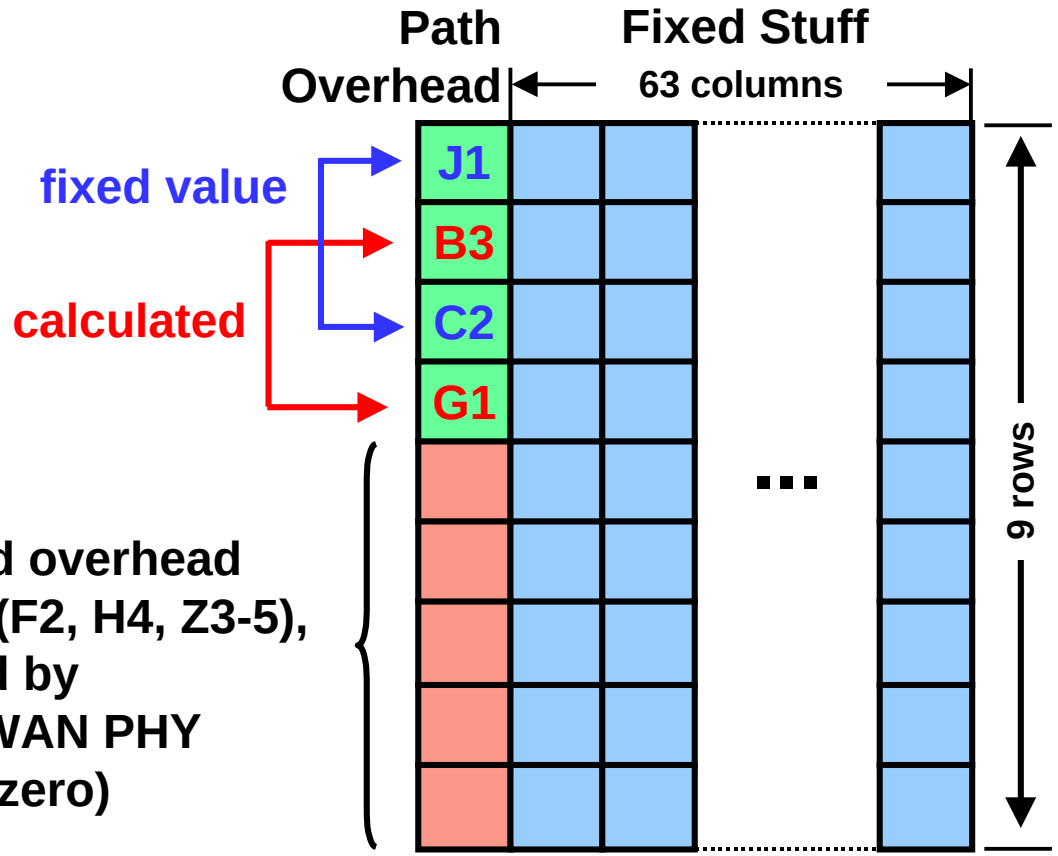


Line Overheads (cont.)

- **K1 and K2**
  - Fixed values: K1 = 00000001, K2 = 00010000
  - K1 and K2 are used on the protection line for automatic protection switching signaling. Above settings indicate a working channel rather than the protection channel.

- **S1**
  - Fixed value: 00001111
  - Indicates quality clock information to receiver. Above setting indicates “don’t use for synchronization”

Path Overhead and “Fixed Stuff”



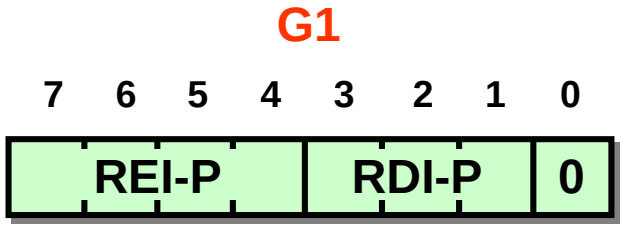
“Fixed Stuff” columns provide compatibility with SONET/SDH byte-interleaving and concatenation rules (set to zero)

Path Overheads

- **J1 (“Path Trace”)**
  - Fixed value: 00000000
- **B3 (“Path BIP-8”)**
  - Used as a Path error monitoring function
  - Calculated value: BIP-8 code (using even parity) over all the octets of the last transmitted SPE before (x^7+x^6+1) scrambling
- **C2 (“Path Signal Label”)**
  - Identifies the contents of the STS SPE (i.e., 10GE WAN PHY)
  - Fixed value: 00011010 (provisional value assigned to 10 GE)

Path Overheads (cont.)

- **G1 (“Path Status”)**
  - Conveys the Path terminating status and performance back to the transmitter (i.e., a PTE)
  - Calculated value:
    - REI-P field = number of bit errors detected with the B3 octet of the last received SPE
    - RDI-P field = Detected defects on the received signal (values are TBD)

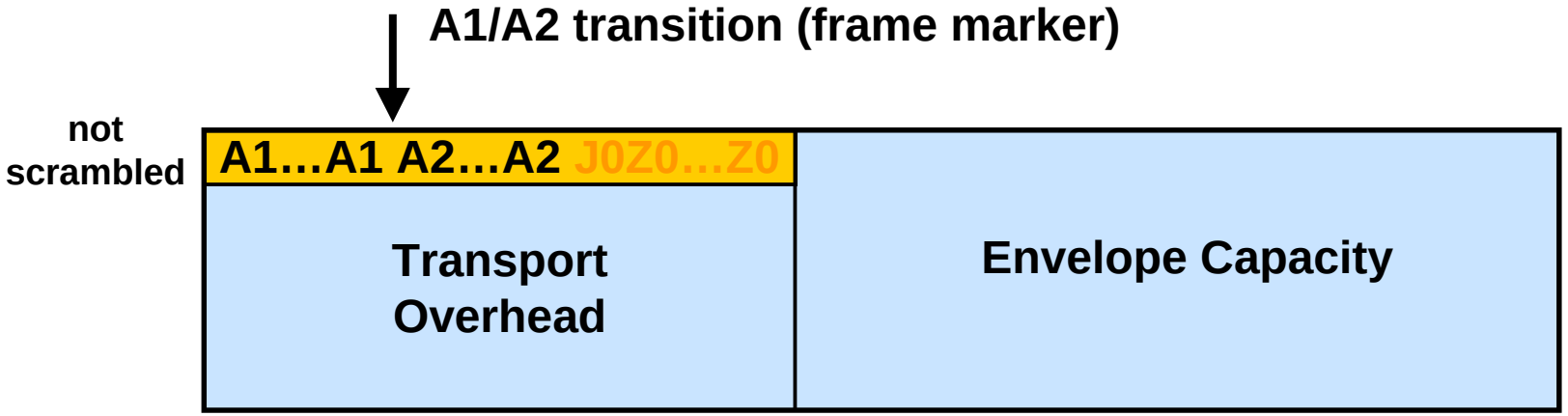


REI-P = Path Remote Error Indication
RDI-P = Path Remote Defect Indication

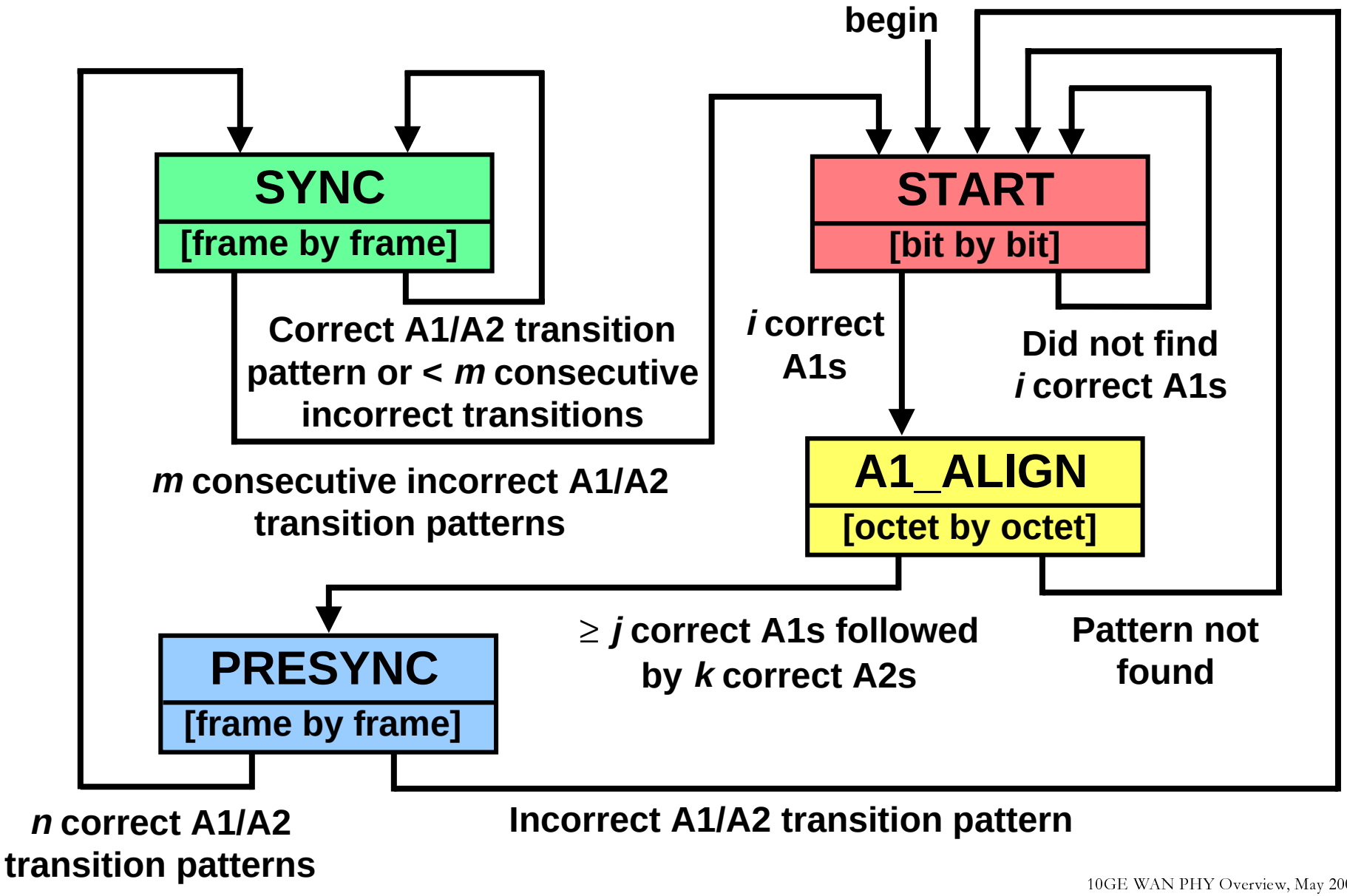
REI-P field
0000 to 1000 = 0 to 8 errors
when received, 1xx1 = 0 errors

WIS Frame Synchronization

- **Uses A1/A2 transition (i.e., frame marker) for frame and octet delineation**
- **Looks for the A1/A2 framing pattern consistently**
  - Expects it to appear once every 155520 octets (length of the frame)
  - When the framing pattern appears in the right place enough times, correct frame synchronization is assumed



Frame Sync Example: State Diagram

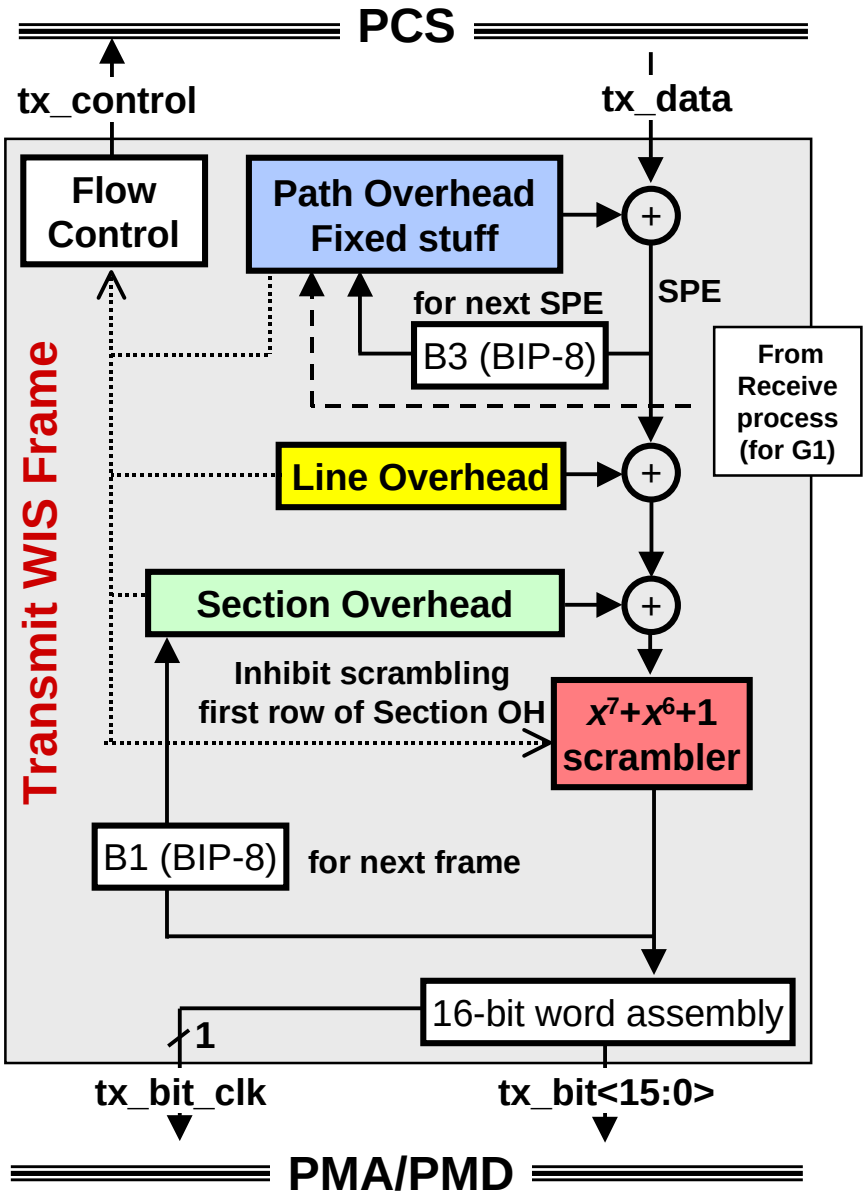


WIS Frame Sync. Performance

- **Example for *m* = 4, A1/A2 transition pattern = 2 A1/A2s**
  - Probability of frame loss $\approx 1.049 \times 10^{6} \times \mathrm{BER}^{4}$
    $= 1.049 \times 10^{-42}$ (@ BER = 10^{-12})
  - Average interval to frame loss
    - $\approx 3.7 \times 10^{30}$ years (@ BER = 10^{-12})
      (> estimated age of observable universe, i.e., ~ 10^{10} years)

- **More robust implementations are possible, e.g., see**
  - **"10GE WAN PHY Delineation Performance"**
  - **http://grouper.ieee.org/groups/802/3/10G_study/public/email_attach/delineation_perf.doc**

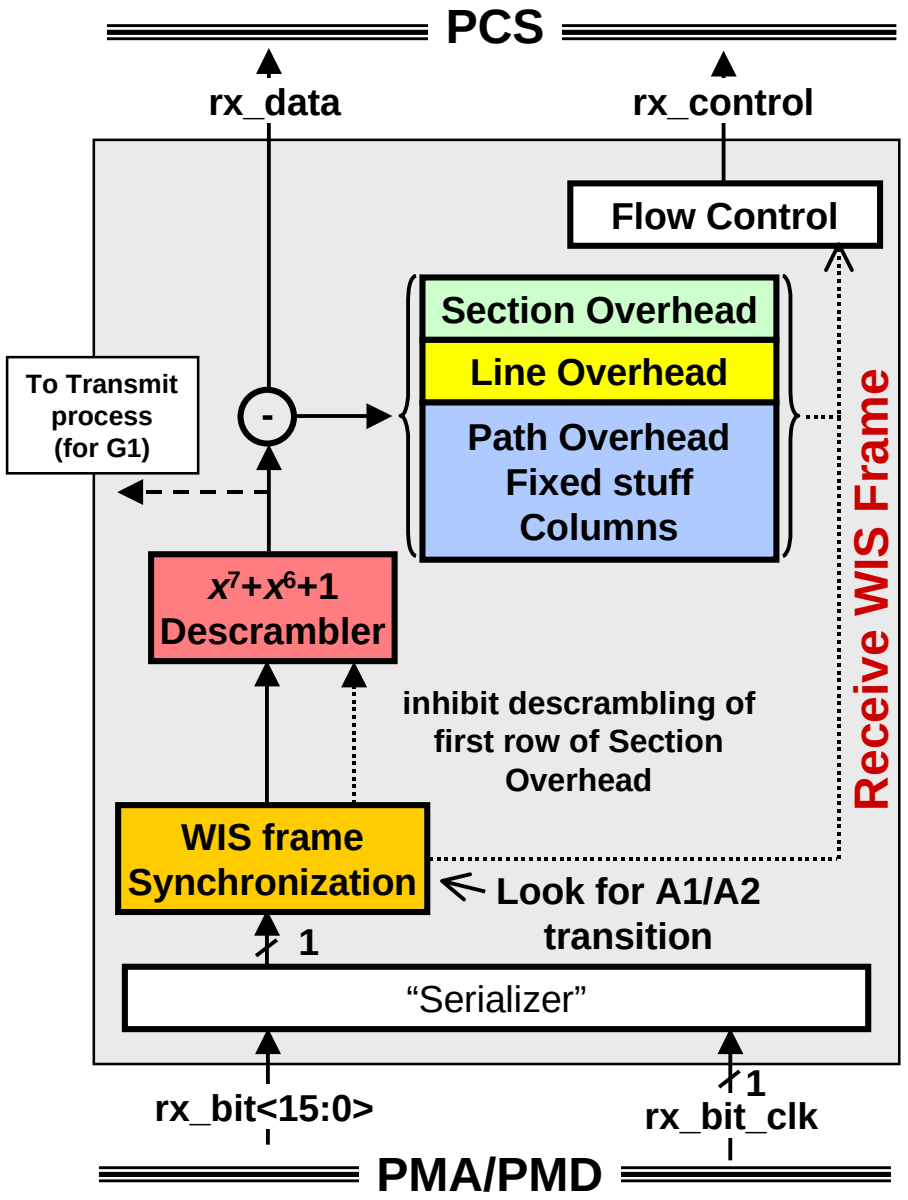
Reference Diagram: Transmit WIS Frame



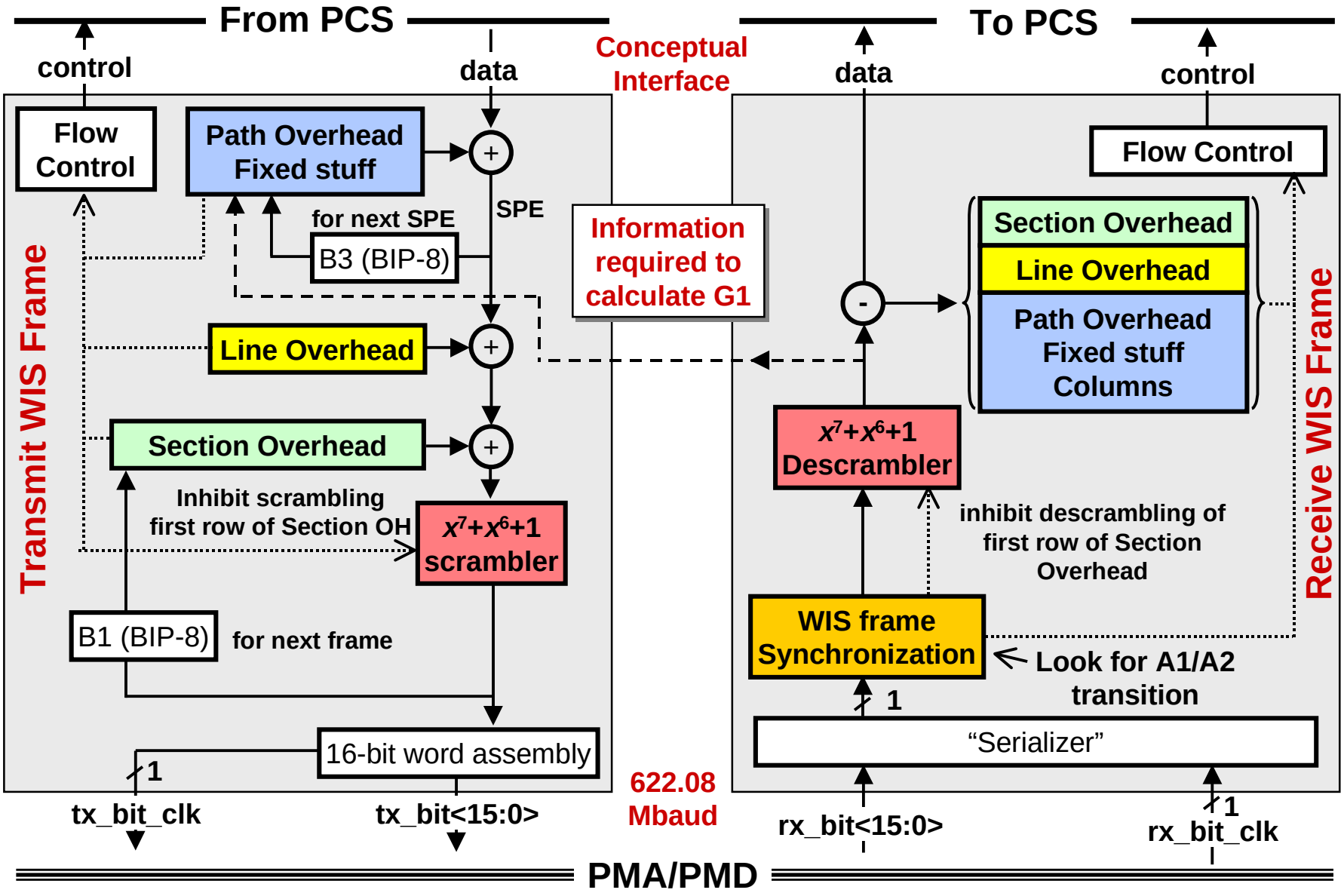
- **Functional View**
- **WIS frame formation (stages)**
  - (1) Path Overhead and fixed stuff columns
  - (2) Line Overhead
  - (3) Section Overhead
  - (4) Scramble with x^7+x^6+1 **(**first row of Section Overhead, i.e., A1/A2, J0, and Z0, is not scrambled)
  - (5) 16-bit words are transmitted to PMA/PMD (for 16-bit Parallel I/F)

Reference Diagram: Receive WIS Frame

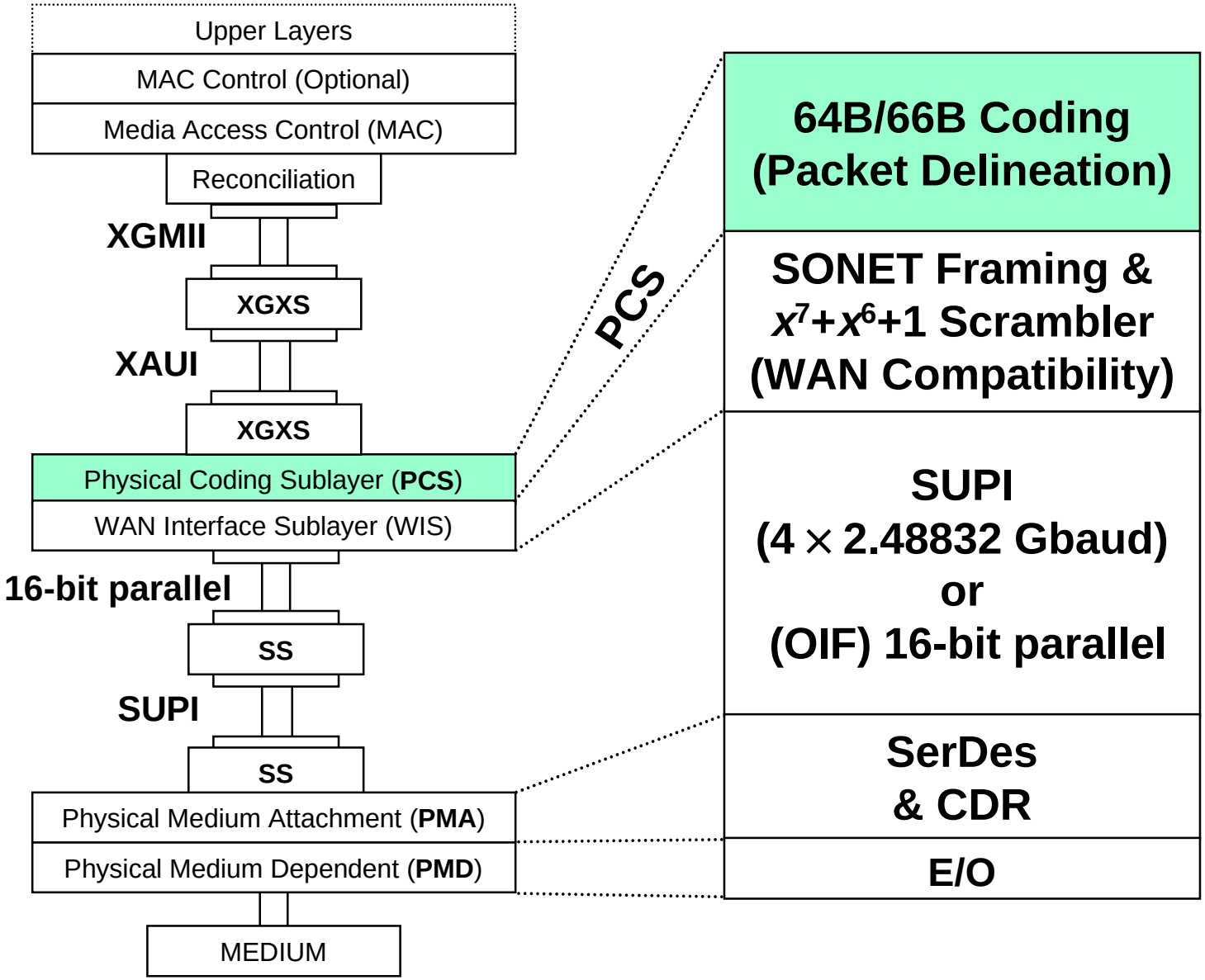
- **Functional View**
- **WIS frame processing (stages)**
  - (1) “Serialize” received signal (figure shows 16-bit Parallel I/F)
  - (2) WIS frame synchronization and octet delineation
  - (3) Descramble with x^7+x^6+1 (first row of Section Overhead is <u>not descrambled)</u>
  - (4) Extract Section Overhead, Line Overhead, Path Overhead, Fixed Stuff columns
  - (5) Remaining octets = payload



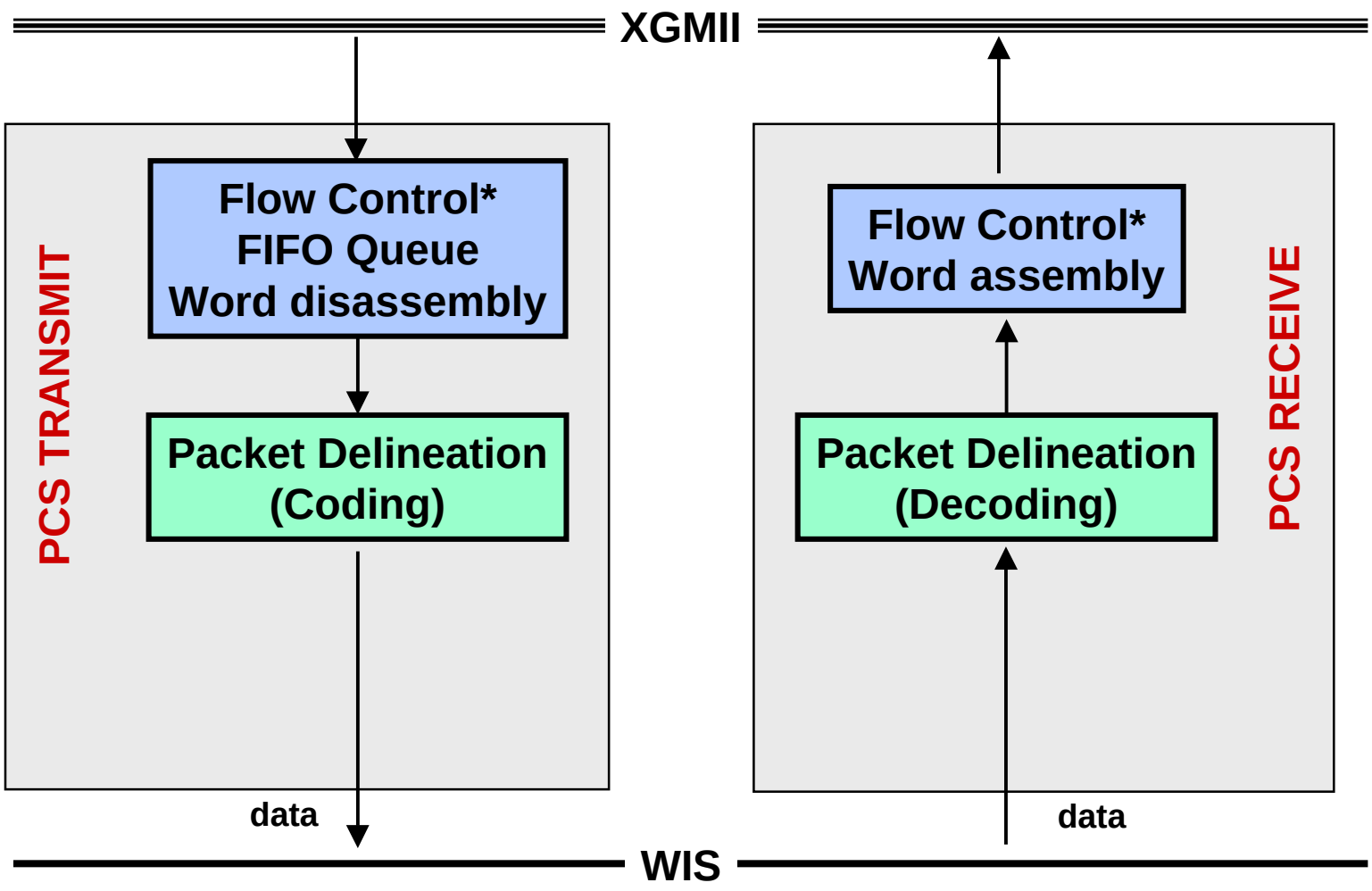
WIS Reference Diagram



Packet Delineation



PCS Reference Diagram



* Required depending on rate control mechanism

64B/66B

- **Data Codewords have “01” sync preamble**

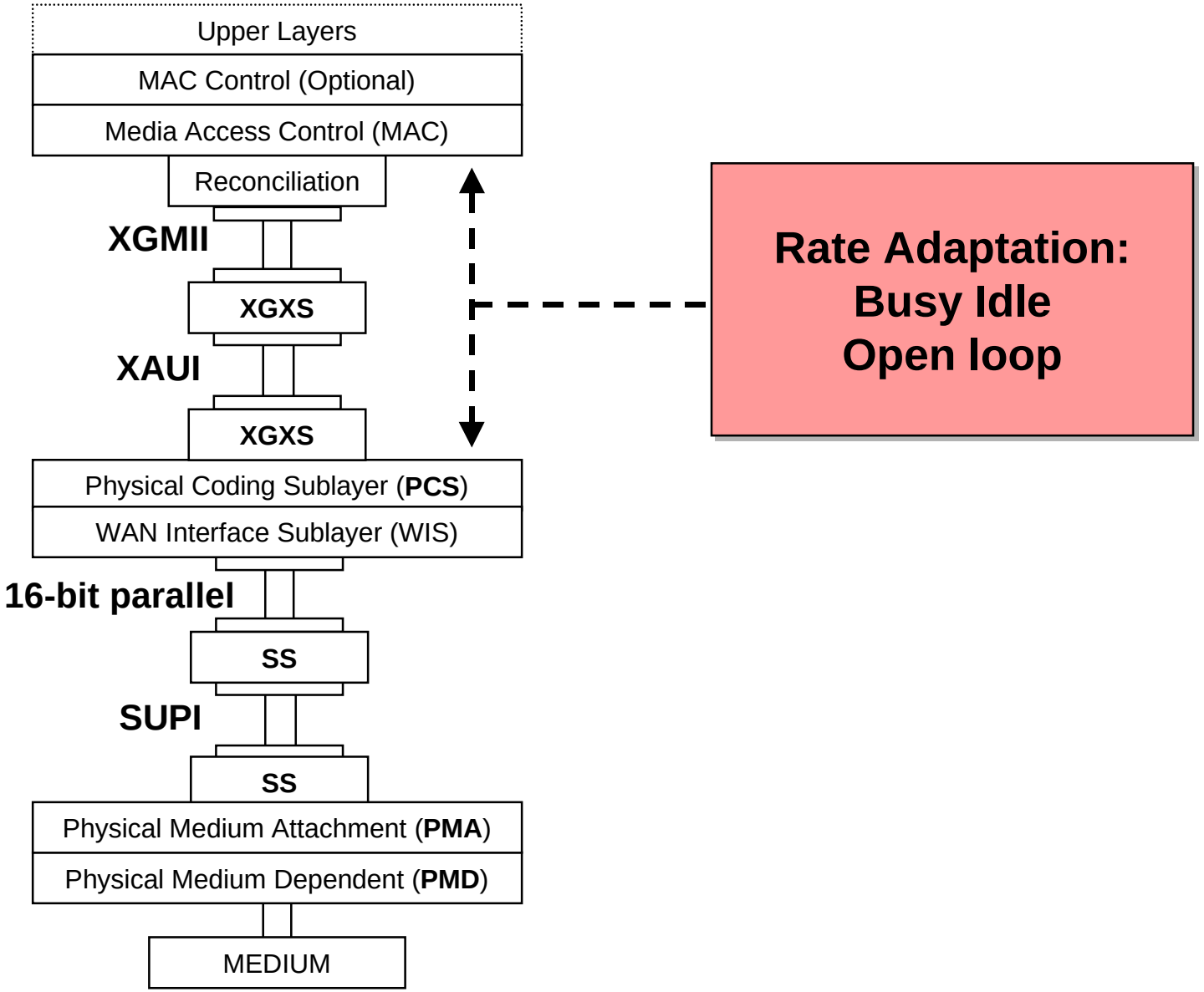
- **Mixed Data/Control frames are identified with a “10” sync preamble**

  - Both the 8-bit type and 56-bit data/control fields are scrambled

- **Uses $x^{58}+x^{19}+1$ self synchronous scrambler**

Reference: http://grouper.ieee.org/groups/802/3/ae/public/mar00/walker_1_0300.pdf

Rate Adaptation



Rate Adaptation

- **Several rate adaptation proposals**
  - Open loop
    - MAC knows the data rate of the PHY and pauses transmissions at frame boundaries
  - Busy Idle
    - PHY sends “Busy Idle” to MAC during IPG
    - MAC pauses transmission at frame boundary

- **10GE WAN PHY will work with any of them**

Summary

- **WIS (WAN Interface Sublayer)**
  - $x^7 + x^6 + 1$ scrambler
  - SONET framing, overheads, and frame synchronization
- **PCS (Packet Delineation Function)**
  - 64B/66B coding
- **Optional PMD Interfaces**
  - 16-bit Parallel (OIF) and SUPI (4×2.48832 Gbaud) for WWDM
- **Rate Adaptation**
  - Several proposals: Open loop, Busy Idle