



Aggregation at the Physical Layer

Howard Frazier
IEEE P802.3ba Task Force
Portland, OR
23-January-2008

APL Goals

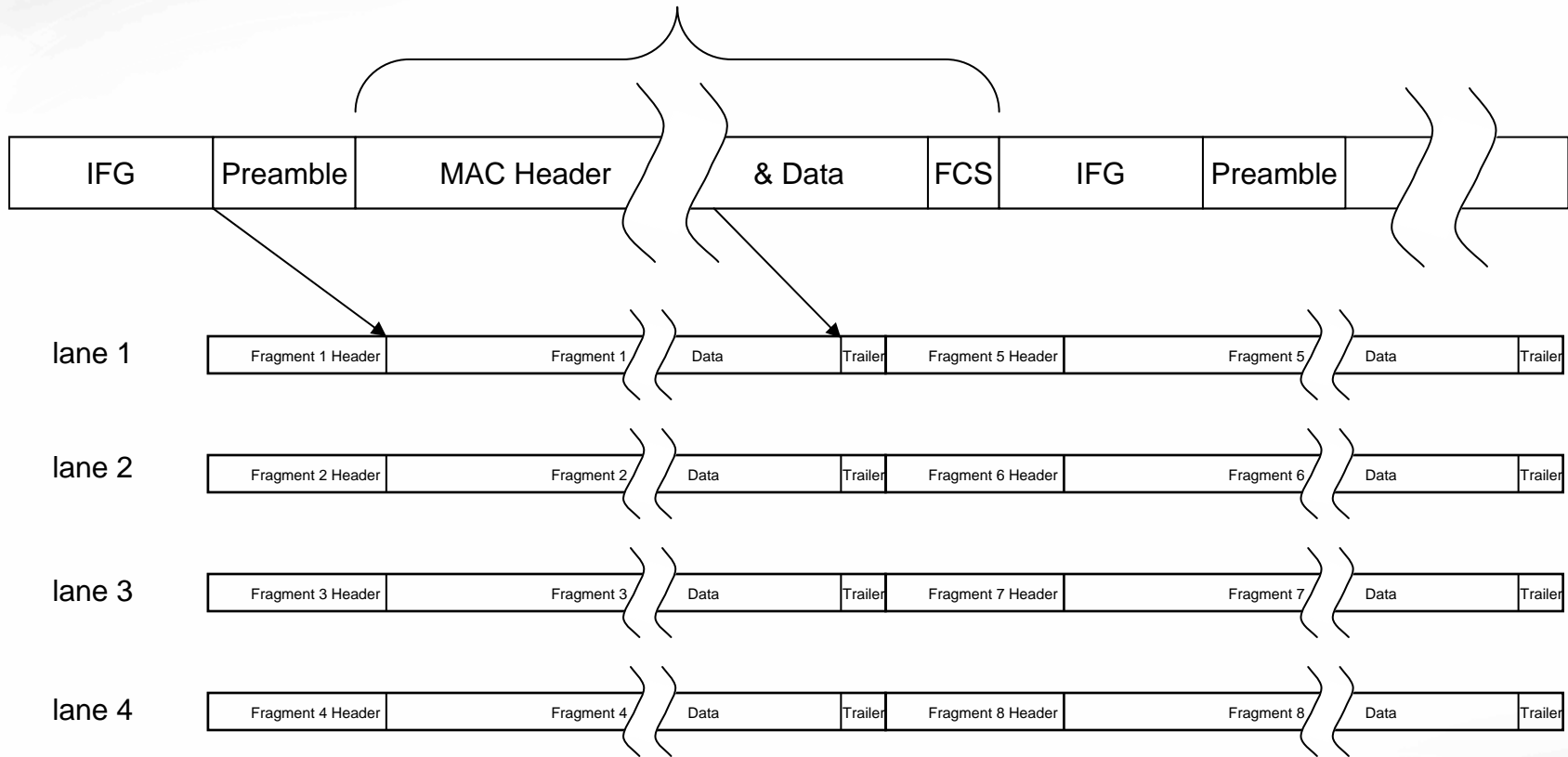
- Define a mechanism that efficiently “bonds” physical links together to achieve a higher aggregate rate
- Support aggregate rates of 40 G and 100 G
- Keep the mechanism coding independent
- Be compatible with all existing 10 G PHYs
- Be compatible with all existing 10 G sublayer interfaces
- Be compatible with all future n G PHYs
- Support lane counts of 1, 2, 4, 5 and 10
- Support lane rates of 10 G, 20 G, 25 G, 50 G and 100 G
- Accommodate any reasonable amount of lane to lane skew
- Preserve temporal ordering of packets

APL Key Concepts

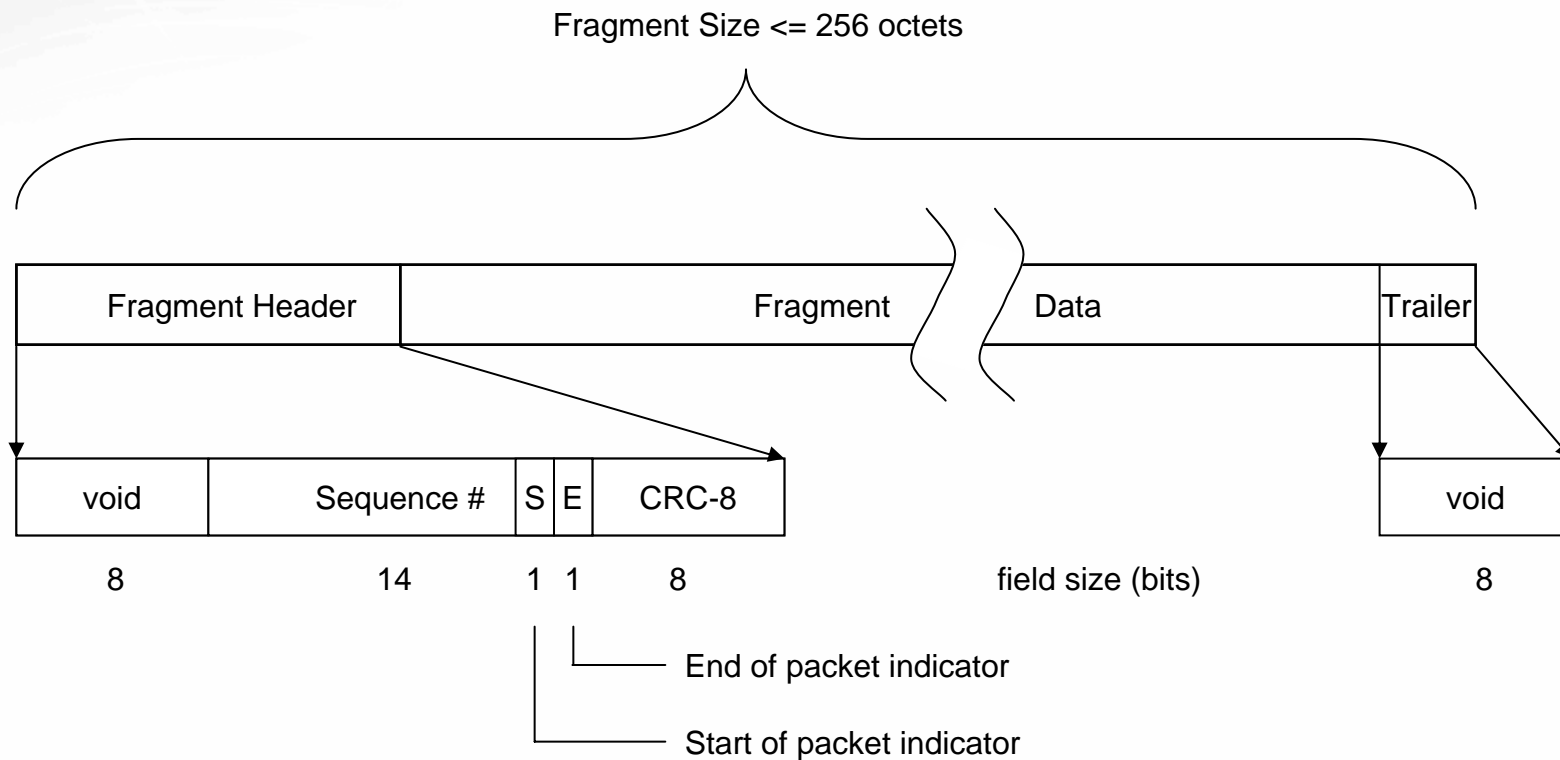
- Transport all of a packet, including the preamble
- Decimate packets into fragments at the transmitter
- Fragments are delimited just like packets on any given interface
- Distribute fragments across the available lanes
- Tag fragments with a sequence number so that temporal order of packets can be preserved
- Reassemble fragments at the receiver

APL Fragmentation

Packet Size \geq 64 octets



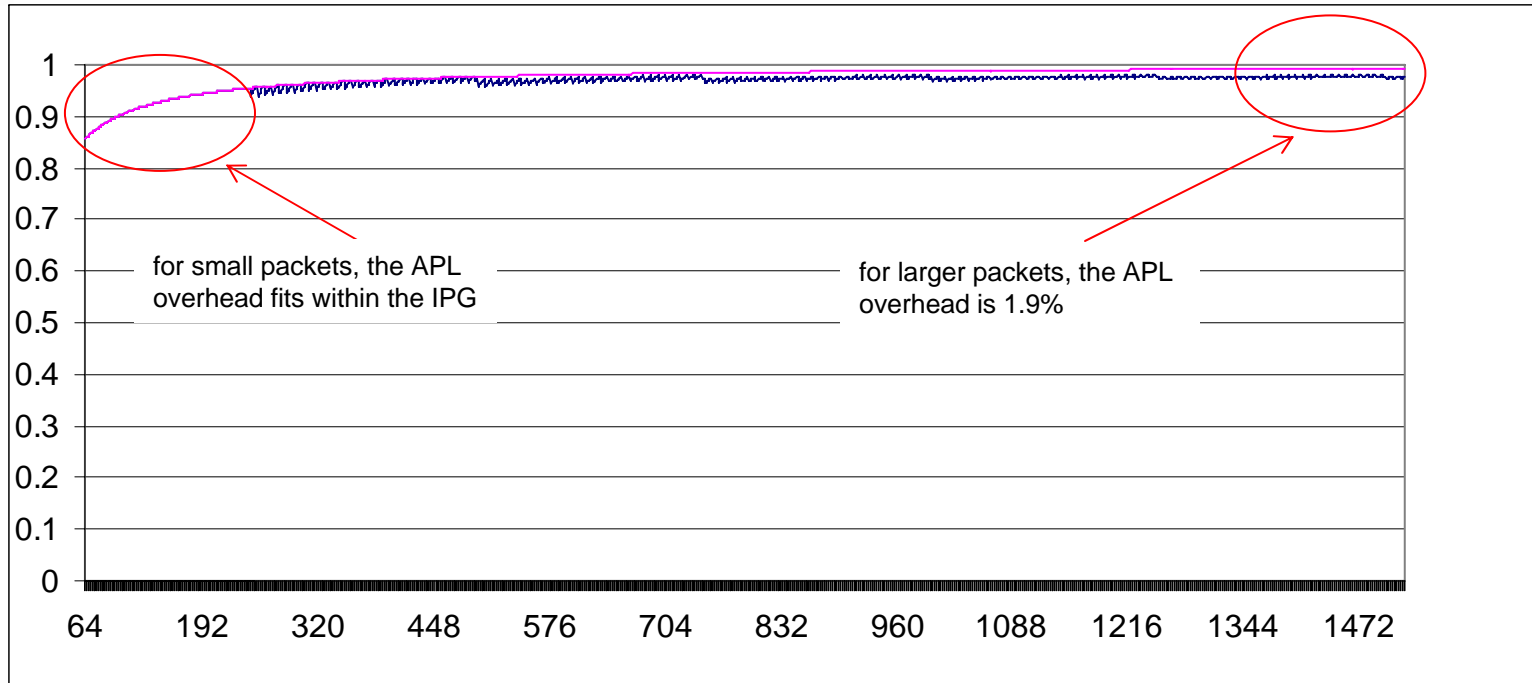
APL Fragment Format



Decimation

- The bulk of a packet is decimated into the largest size (256 octet) fragments
- The tail of a packet is decimated into a fragment with a maximum size of 256 octets, a minimum size of 16 octets, that is a multiple of 8 octets in length
- For certain packet sizes (those which produce a remainder after decimation of 244, 245, or 246 octets) the transport efficiency can be improved a little bit if the penultimate fragment size is 248 octets, and the last fragment size is 16 octets, rather than 256 octets and 16 octets
 - 8 gates required to detect these cases

APL Transport Efficiency



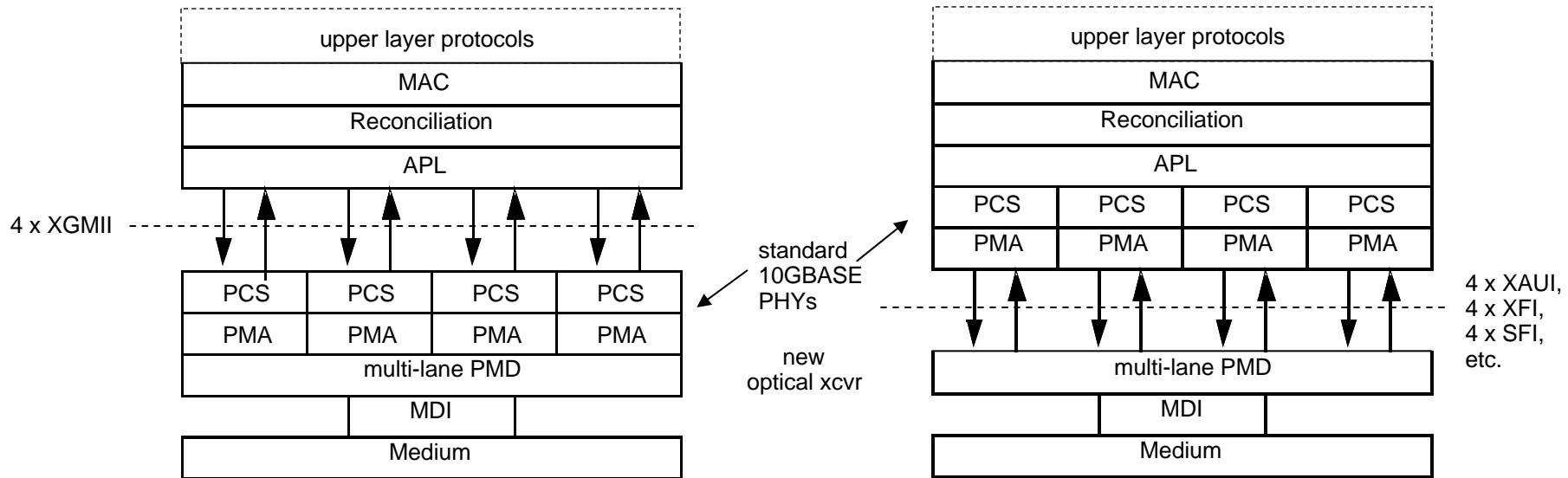
Adapting the MAC rate

- The ifsStretch mechanism is already defined in Clause 4
- The MAC rate can be reduced slightly to account for APL overhead
- Increase the interframe spacing by 5 octets for every 250 octets of a packet

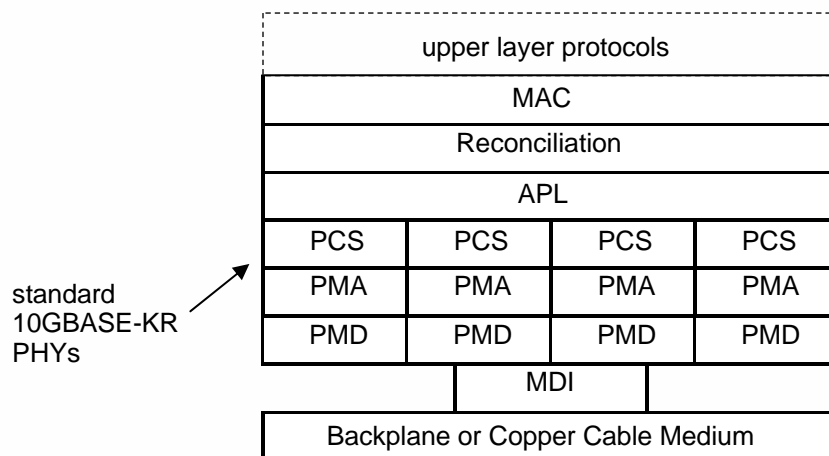
Inserting Idles

- "Idles" must be inserted into the data stream on a periodic basis to allow for clock tolerance compensation
- The APL sublayer maintains a count of the number of octets transmitted on each lane – one counter per lane
- A counter is reset whenever 8 octets of "Idle" are transmitted on the corresponding lane
- The APL sublayer inserts 8 octets of "Idle" (no data transmission) at the next fragment boundary after 8192 octets of data have been transmitted
- Overhead = $8/8192 \approx 0.1\%$

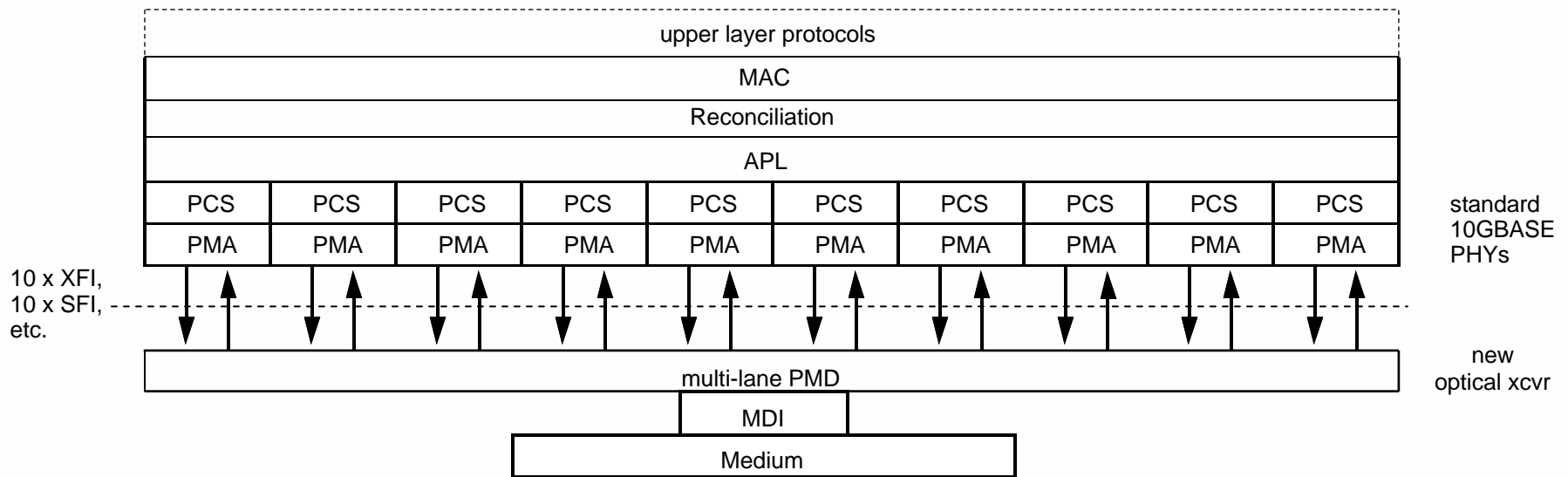
Using APL to achieve 40 G



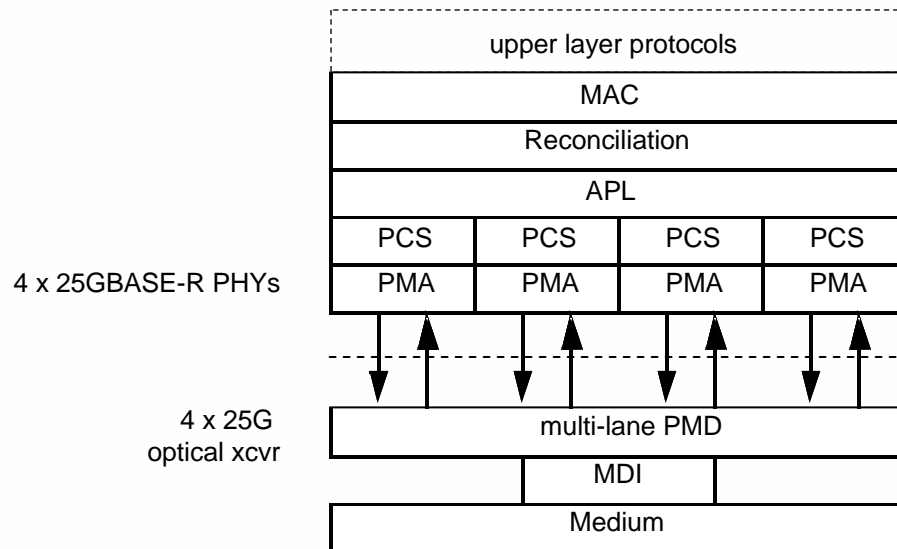
Using APL to achieve 40 G



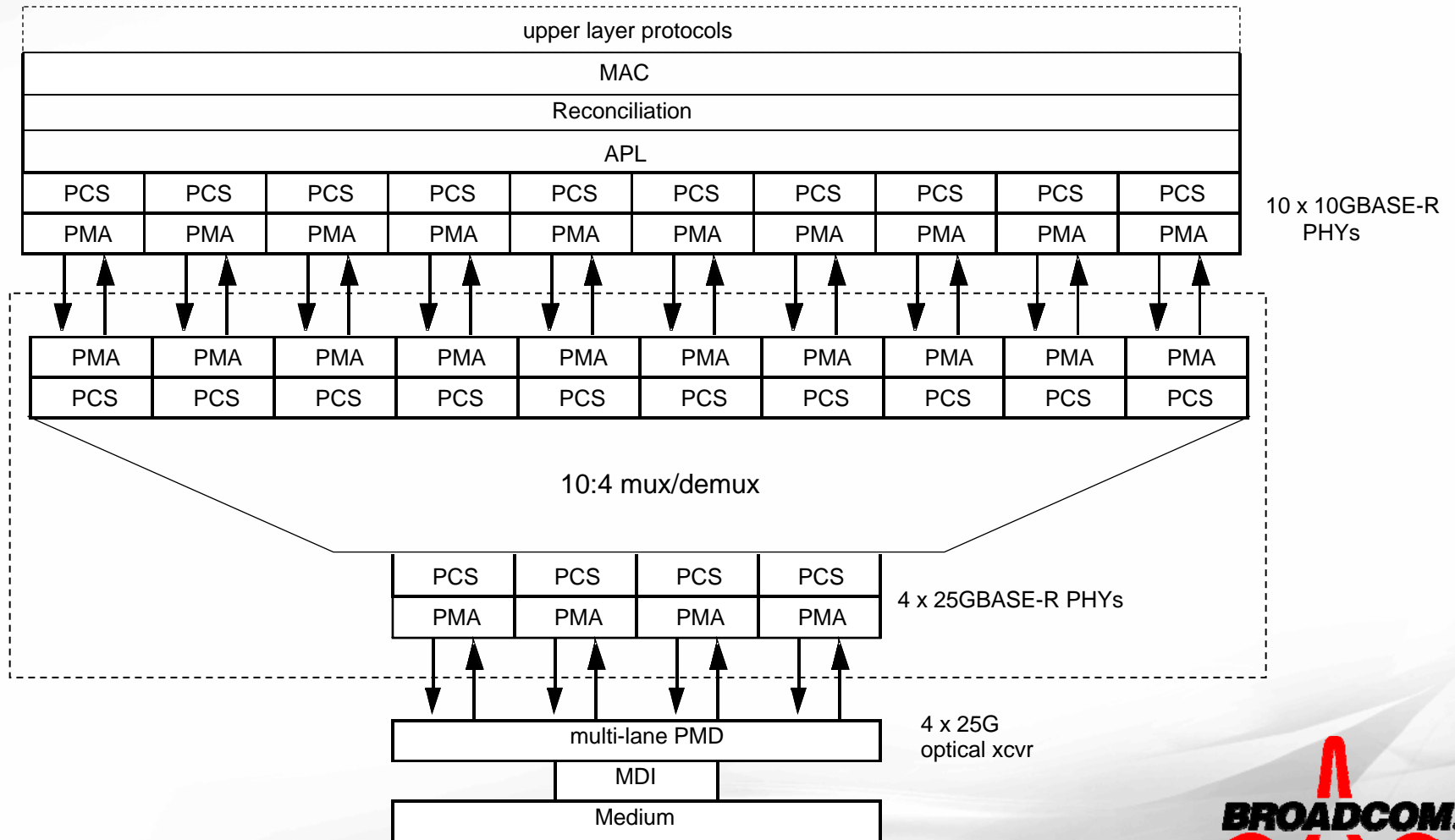
Using APL to achieve 100 G (10 x 10)



Using APL to achieve 100 G (4 x 25)



10 x 10 into 4 x 25



APL Summary

- Efficient bonding mechanism that supports 40 G and 100 G aggregation of lower speed PHYs
- Compatible with all existing 10 G PHYs and sublayer interfaces
- Future proof mechanism that can be used with any future physical layer (e.g. 4 x 100 G, 10 x 100 G)