CTBI:
A simple lane bonding mechanism for both 40GE and 100GE Interfaces

Gary Nicholl

Cisco

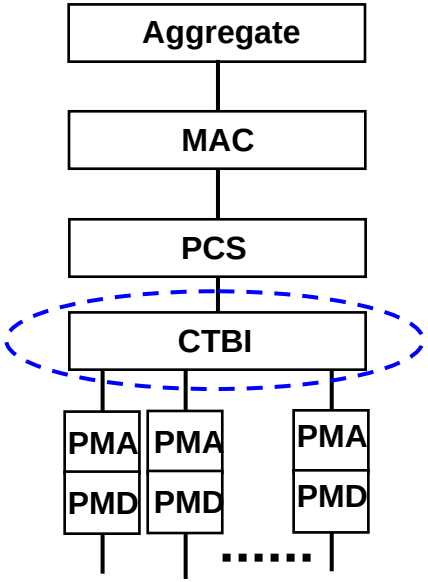
Topics

- **Lane Bonding for 40GE/100GE**
- **CTBI Refresher (Goals and Key Concepts)**
- **Implementation**
  - Considerations for host board and optical modules
  - 40GE implementation examples
  - 100GE implementation examples
  - Comparison of CTBI and APL
- **Summary**

Lane Bonding for 40GE and 100GE

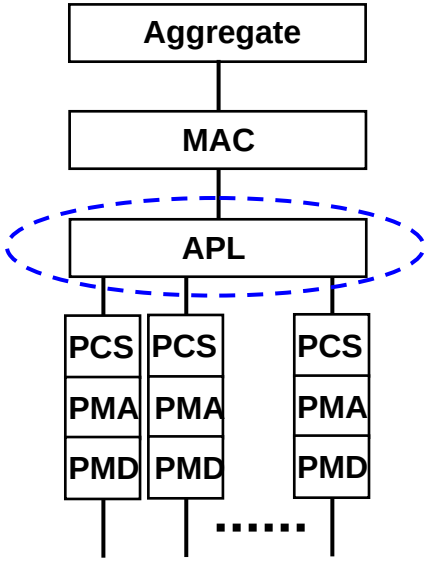
- **40G and 100G interfaces will use (require) multi-lane PHYs**
  - lanes could be cables, fibers or wavelengths
- **Need a mechanism to effectively 'bond' the physical lanes together**
- **There are several proposals for doing this:**
  - CTBI, APL, PBL
- **At a certain level all these schemes are all identical.**
  - Tx: stripe aggregate data across multiple low speed links
  - Rx: collect data from multiple low speed links, and reconstruct aggregate
- **The differences are both in exactly where in the protocol stack the striping is performed, and in how it is implemented.**

Lane Bonding Options for 40GE and 100GE



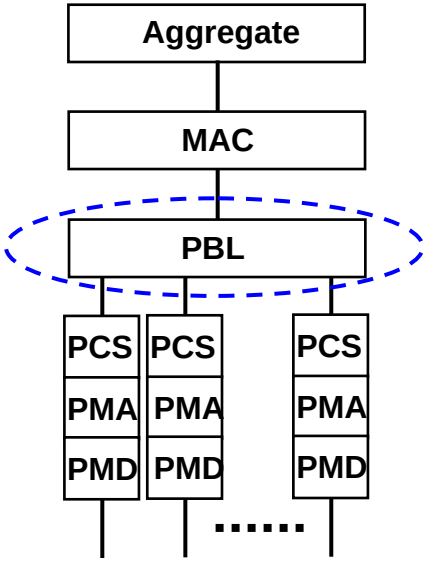
CTBI

- word (8 byte) striping below PCS
- one MAC and one PCS per interface
- can be extended using virtual lanes, to support simple optical modules at all rates and reaches



APL

- frame striping above PCS
- frames decimated into variable length fragments (SAR), tagged and distributed across lanes
- one MAC, but multiple PCSs (one per PMD lane)
- potential to reuse existing 10G PHY devices



PBL

- word (8 byte) striping above PCS
- one MAC, but multiple PCSs (one per PMD lane)
- potential to reuse existing 10G PHY devices

CTBI Goals

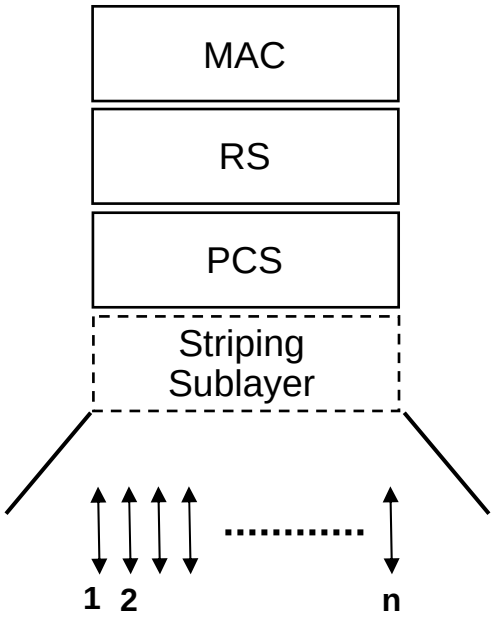
- **Define a mechanism which effectively bonds physical lanes together to achieve a higher aggregate data rate**
- **Support aggregate rates of both 40G and 100G**
- **Enable simple optical modules (for all rates and all PMD types)**
  - PMA/PMD only
  - keep protocol stuff on the host board (where it belongs**)**
- **Decouple electrical and optical lane widths**
  - allow both to evolve independently (get narrower and faster), with minimal design churn. Ideally with no changes to upper layers (MAC,PCS,etc)
- **Reuse of existing 802.3 clauses where possible**
  - 64B66B PCS (Clause 49), 10G-BASE-KR FEC (Clause 71), etc
- **Minimize the number and complexity of any new clauses**
- **Support electrical lane widths of 10, 4, 2, and 1**
- **Support optical lanes widths of 10, 5, 4, 2 and 1**
- **Accommodate any reasonable amount of lane to lane skew**

CTBI Key Concepts

- One MAC and one PCS (irrespective of # of PMD lanes)
- Both MAC and PCS run at the aggregate rate (40G or 100G)
- Decimate output of PCS into 66-bit words (easy)
- Distribute the 66-bit words across n (virtual) lanes

  40G: n=4, 100G: n=20
- Tag each (virtual) lane with a unique marker (special 66-bit word inserted on a periodic basis).
- Use tags to identify, align and reorder (virtual) lanes at the receiver
- Reassemble aggregate signal and handoff to PCS
- Virtual lanes are mapped to both electrical and optical (PMD) lanes, through very simple bit muxing functions

CTBI Key Concepts – In Pictures



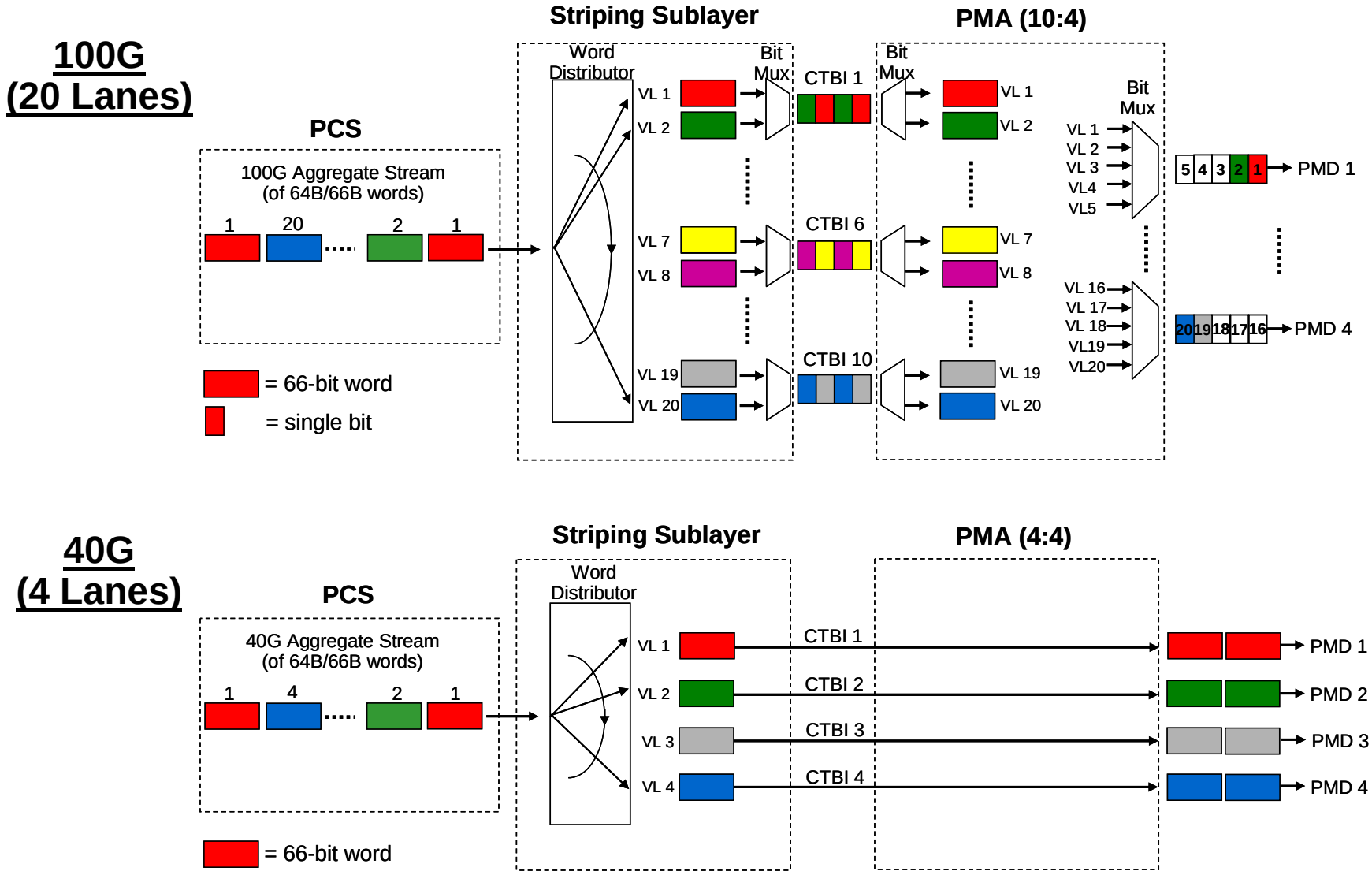
Clause 4

Clause 49 (64B66B PCS)

New Clause

- **Lane Striping**
- **Instead of physical lanes, it's virtual lanes (nothing scary !)**
- **Virtual lanes mapped to elect and optical lanes, with simple bit muxing**
- **Virtual Lanes offer 2 key advantages:**
  - **decouples electrical and optical lane widths**
  - **enables very simple bit mux PMAs (for all PMD flavors)**

CTBI Key Concepts – In Technicolor !!

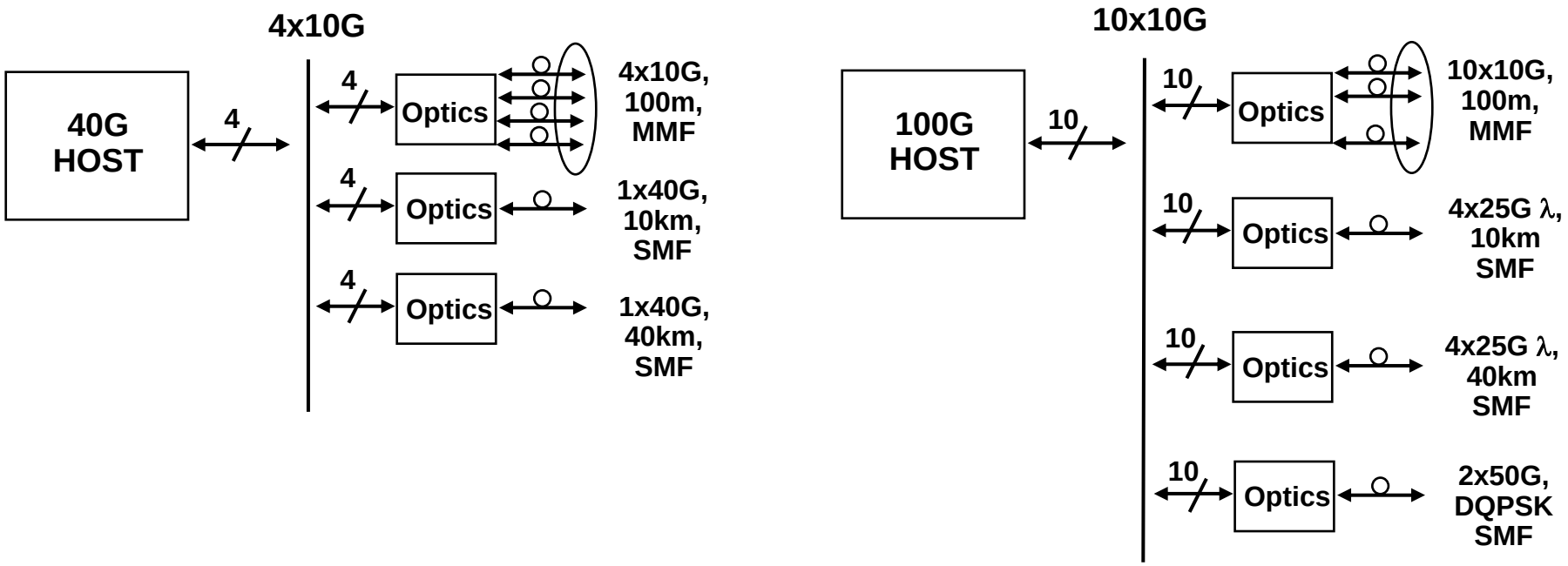


Implementation Considerations

- **Common Host Board + Pluggable modules for each PMD type.**
- **Host Board Considerations:**
  - single vendor environment > complexity not as critical
- **Optical Module Considerations:**
  - pluggable, multi-vendor environment > keep it simple !!
  - historically optical modules are simple PMA/PMD devices and protocol agnostic (bits in, photons out). Examples include GBIC,SFP,XFP,SFP+,300pin,200pin, etc ……
  - the one exception is XENPAK,X2 for 10GE. In this case the XAUI interface drove the PCS block (typically a host board function) down into the module, significantly complicating the design.
  - if we have a choice, would prefer not to repeat the XENPAK experience at 40G/100G

More Optical Module Considerations

- **Common host board drives common elect i/f for all PMDs**
  - ➢ cannot optimize electrical lane width for each specific PMD



Need to decouple electrical and optical lane widths

> BUT still keep module simple !!

Legends

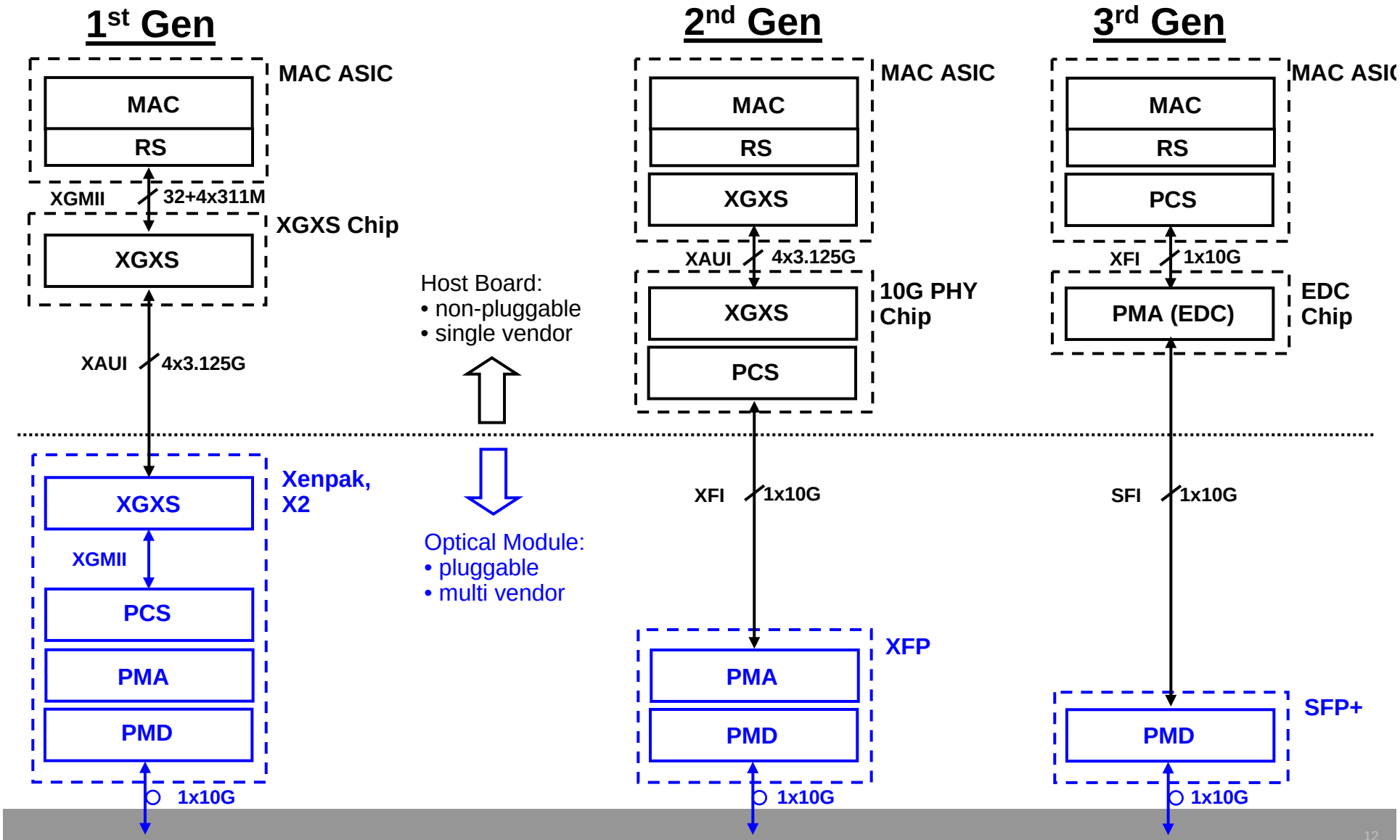
Black = Host Board

Blue = Optical Module

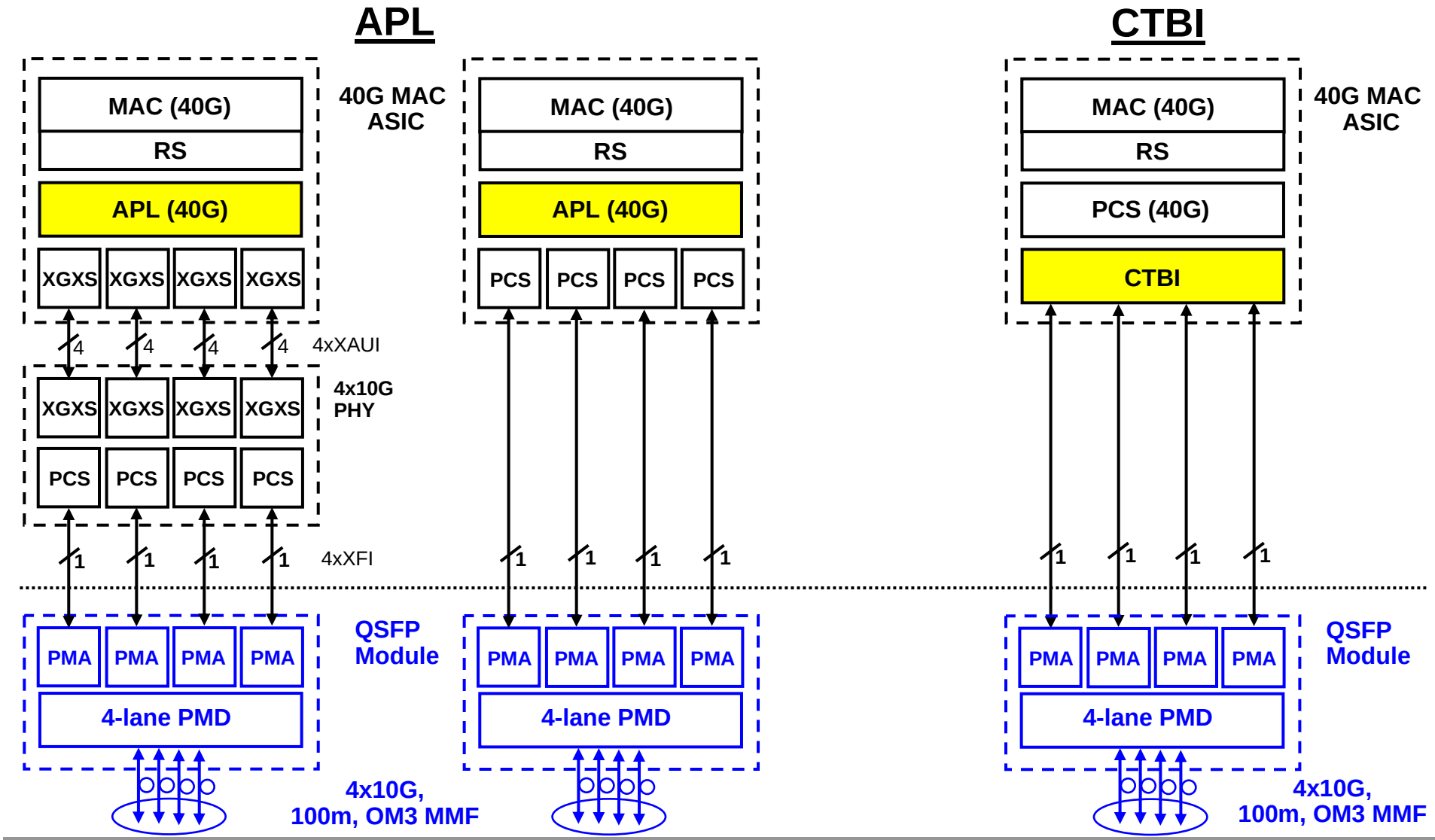
 **= ASIC/FPGA/Chip on host board**

 **= Pluggable Optical Module**

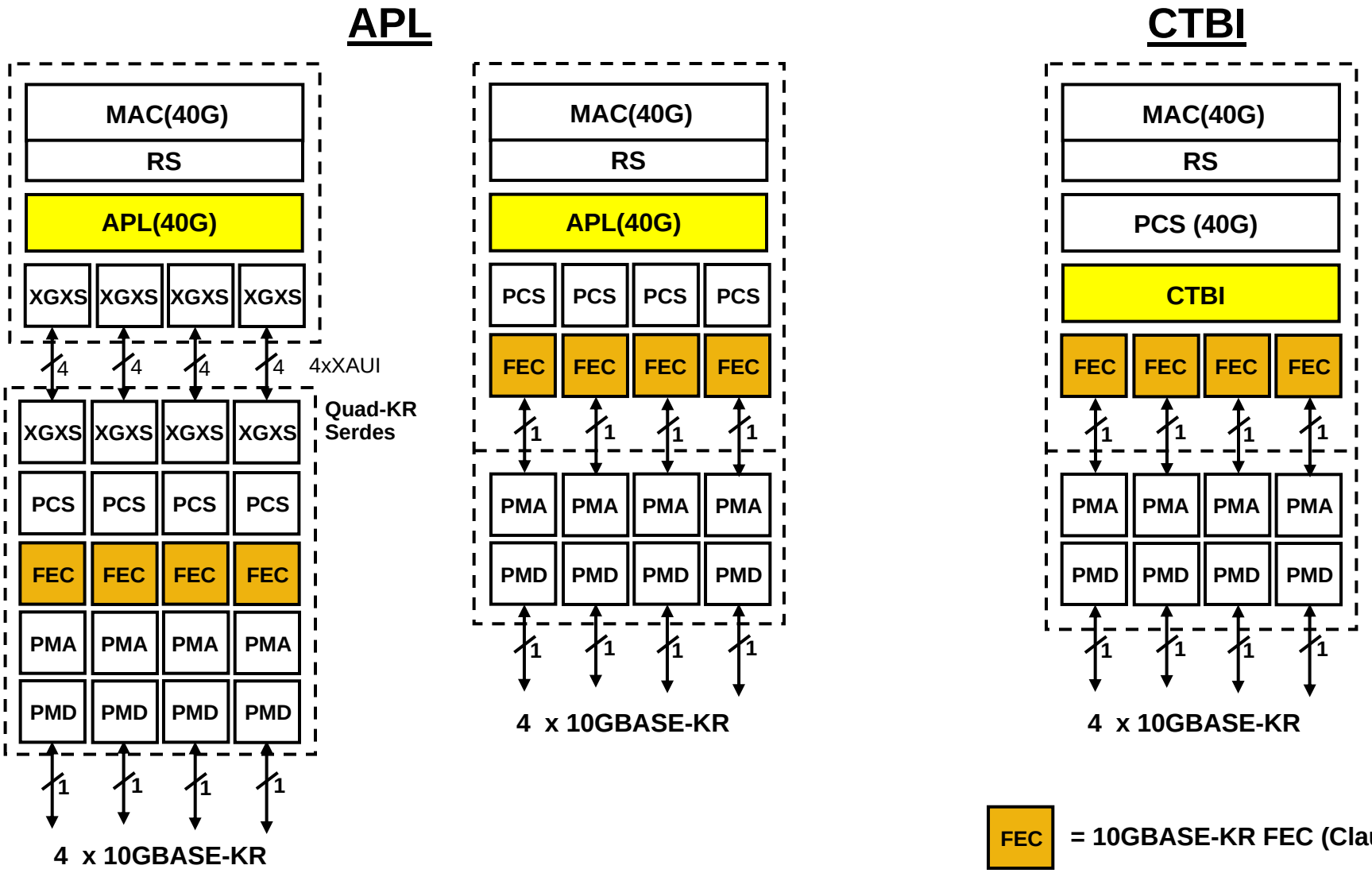
10G Implementations



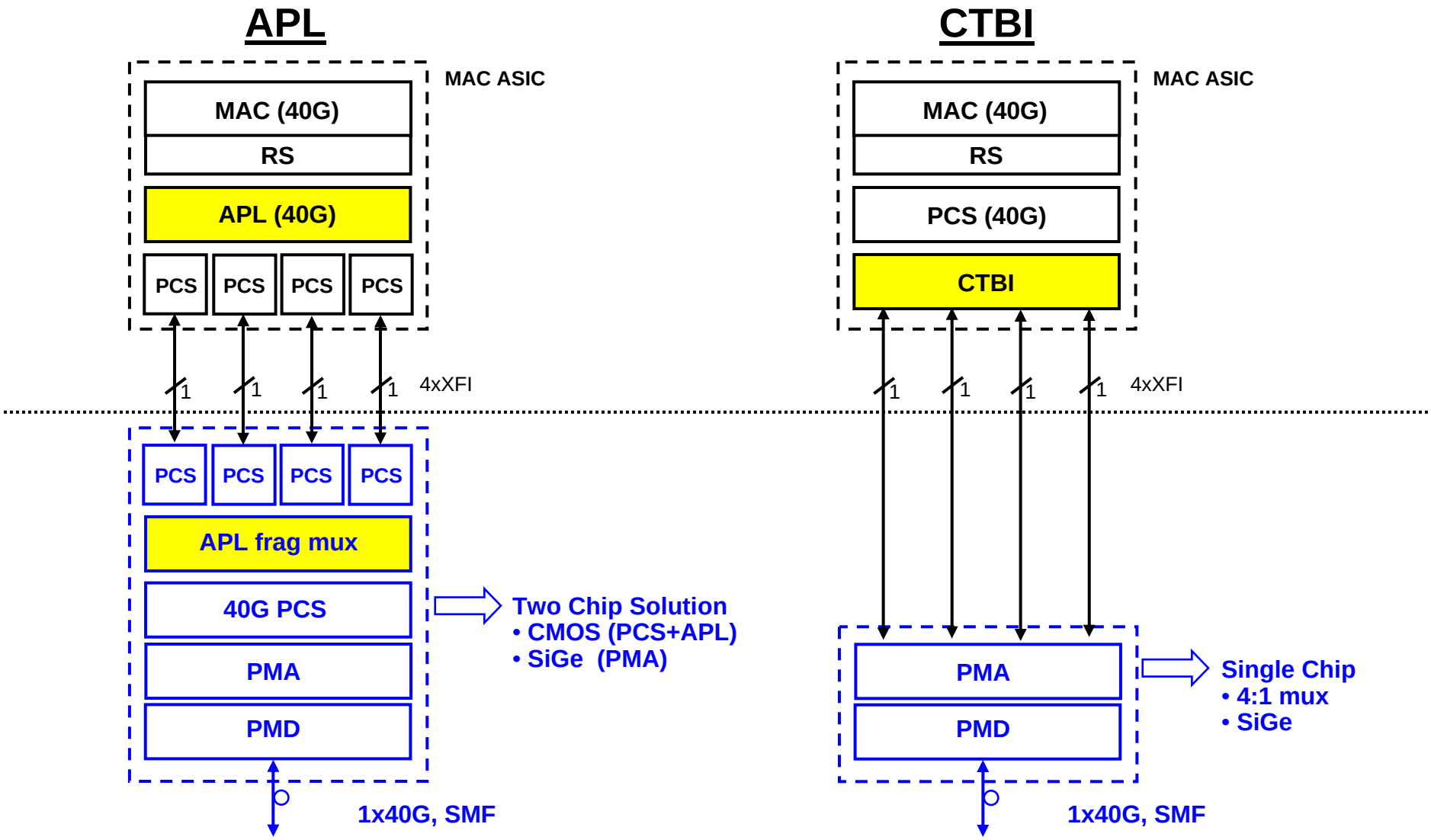
40G - Optical 100m MMF



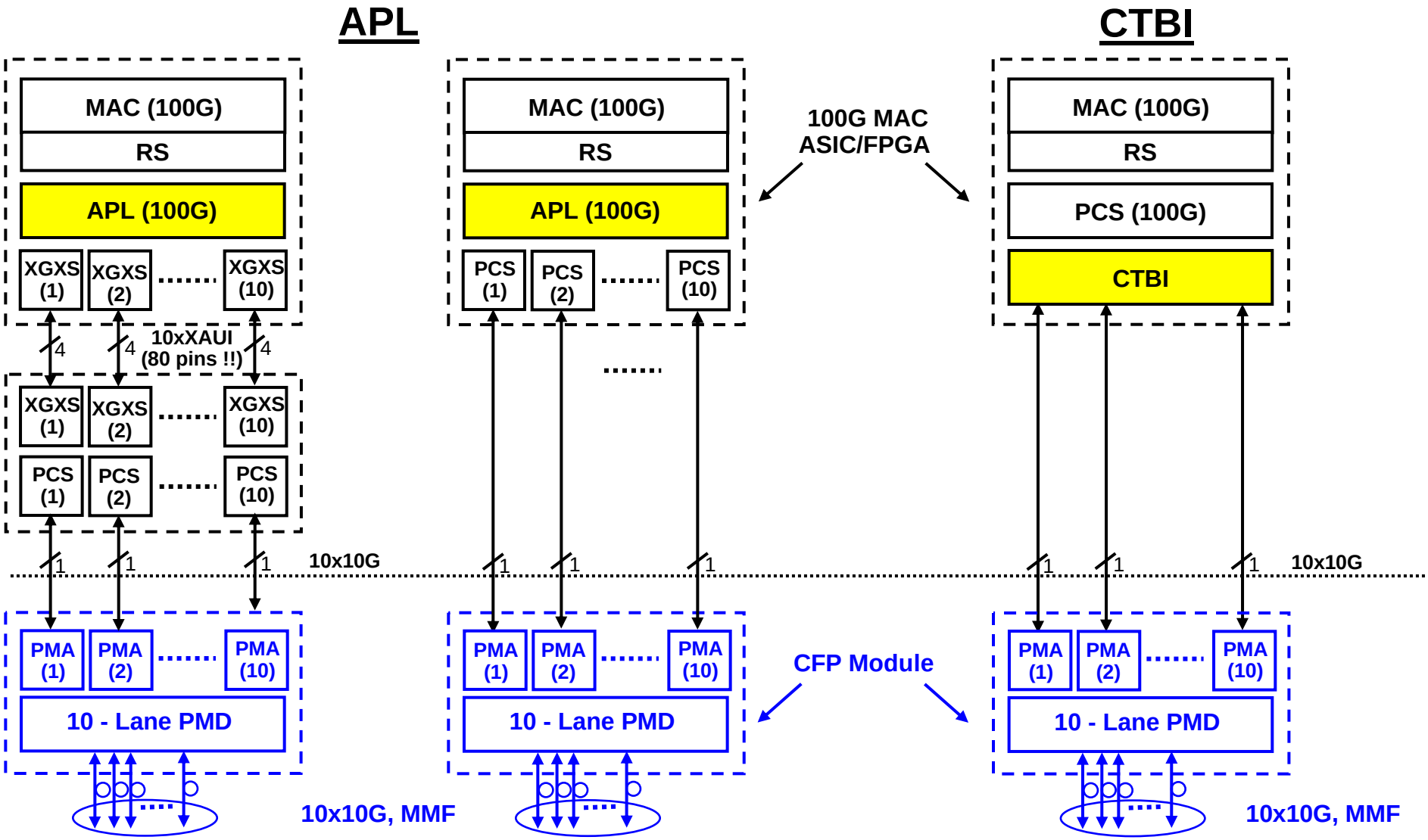
40G - Backplane



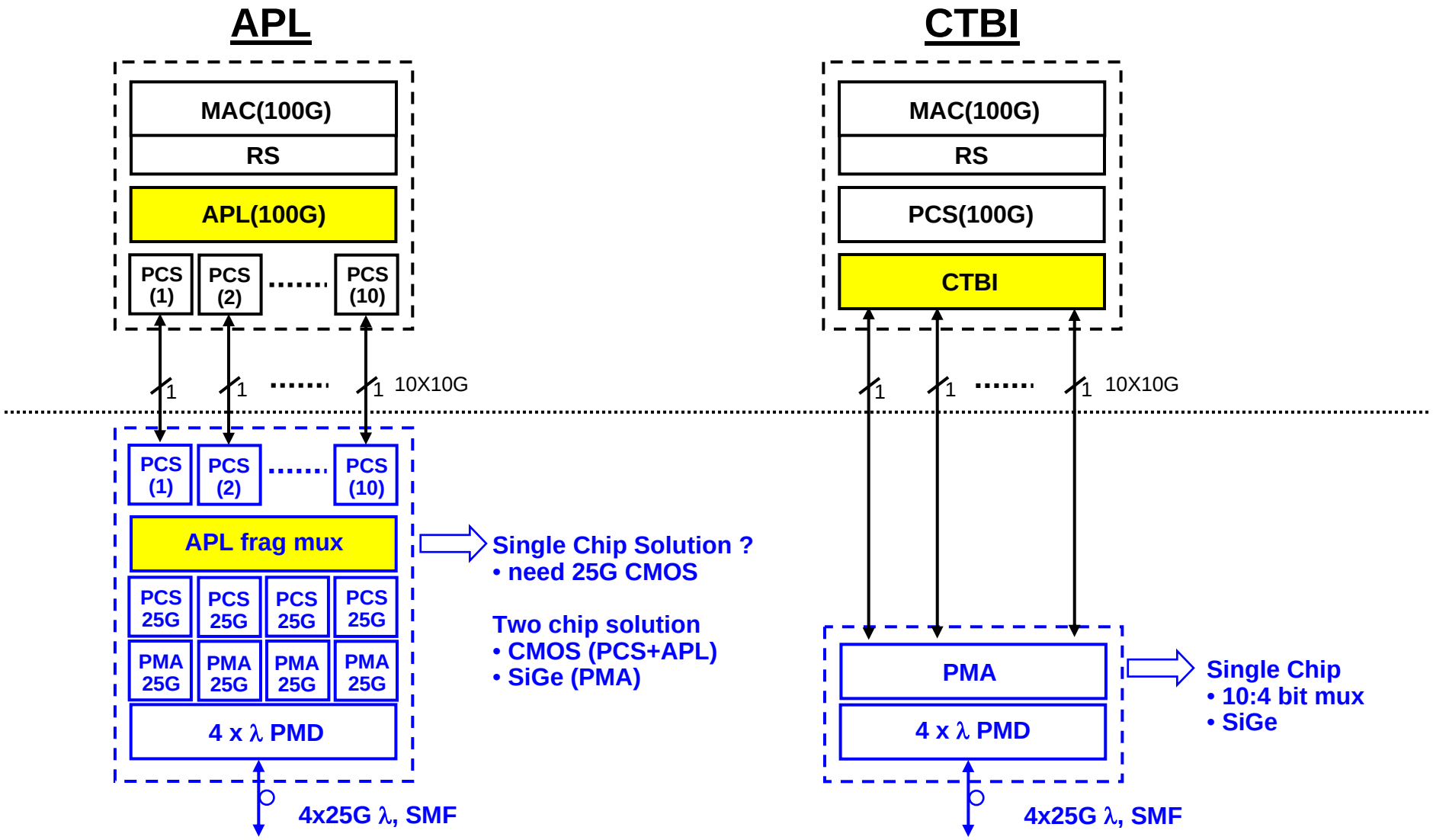
40G - Optical Serial (Future)



100G - Optical 100m MMF



100G - Optical 10/40km SMF



Implementation Comparison (CTBI ‘v’ APL)

- **Host Board:**
  - **at a silicon level it is probably “a wash”.**
  - **at s/w level CTBI may have an advantage due to single PCS**
    - also PCS does also not change as we evolve electrical/optical lane widths
- **40G PMDs:**
  - **no significant difference for MMF PMD (4x10G), or backplane (4xKR)**
  - **APL drives a much more complex serial SMF PMD (Future)**
    - drives protocol functionality down into module (not desired)
    - will require a 2 chip serdes development (CMOS + SiGe)
- **100G PMDs:**
  - **no difference for MMF PMD (10x10G)**
  - **APL drives a much more complex solution for 10km/40km SMF PMD**
    - drives protocol functionality down into module (not desired)
    - likely requires 2 x 100G serdes chip developments (CMOS + SiGe)
    - a two chip solution doubles the high speed I/O (increased power)

CTBI Summary

- **Simple Implementation**
  - for both 40G and 100G applications
  - for all PMDs types (backplane, MMF, SMF, etc)
  - enables simple ‘PMA/PMD only’ optical modules (at all rates and reaches)
  - appears as a single interface to the user (i.e. one MAC and one PCS / interface)
- **Reuse of existing IEEE Clauses**
  - 64B66B PCS (Clause 49)
  - 10G-BASE-KR FEC (Clause 71)
  - just need to write new ‘lane striping’ clause, in addition to new PMA/PDM clauses
- **Scalable and Future Proof**
  - allows electrical and optical lane widths to evolve, without touching the PCS
  - same mechanism can be used for next generation of interfaces (400GE, ITE, etc)
- **Stable proposal**
  - first presented in November 2006 (gustlin_01_1106.pdf)
  - no significant change to goals, concepts or implementation since
  - but we do need a new name !!!