50GbE and NG 100GbE
Technical Considerations for FEC Lane Distribution

Gary Nicholl - Cisco

IEEE 802.3cd Task Force, May 23-25, Whistler - Canada

Motivation

- nicholl_3cd_01_0516 represents a long term optimized solution but doesn't cover some of the more near term technology building blocks that implementers may wish to utilize.
  - e.g. 25 Gb/s IO after the FEC sublayer
- This presentation explores how to modify nicholl_3cd_01_0516 with these considerations and what advantages and disadvantages exist.

Introduction

- nicholl_3cd_01_0516 proposed an architecture where a single RS-544 FEC is distributed over 50Gb/s FEC lanes (single lane for 50GbE, two lanes for 100GbE), in order to match the 50Gb/s lane rate PMD objectives
- There has been discussion around distributing a single RS-544 FEC over 25Gb/s FEC lanes instead (two lanes for 50GbE, four lanes for 100GbE), to allow the FEC to be carried over both 50 Gb/s and 25 Gb/s lane rate AUIs
  - This requires an additional muxing (symbol or bit) stage downstream of the FEC to match a 50Gb/s lane rate AUI (if implemented) and the 50Gb/s lane rate PMD objectives.
- This presentation will address the technical considerations and implications of these two muxing approaches
- A 50GbE use case will be used for the analysis, as 100GbE is essentially identical.

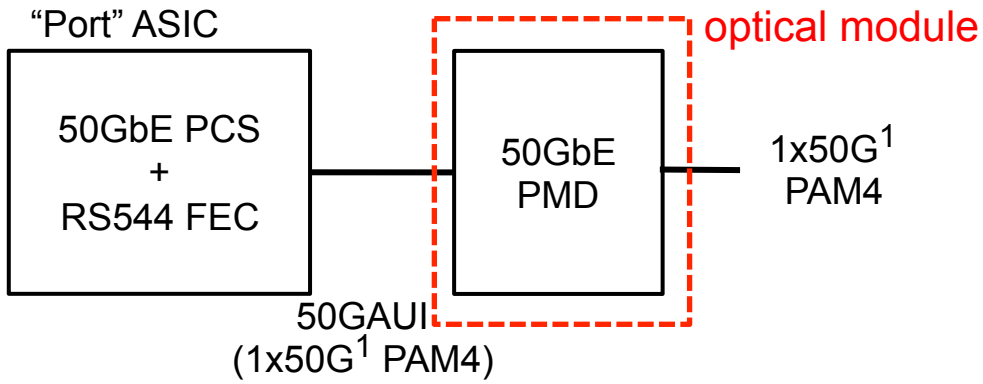
Terminology

Interleaved FEC	In some cases FEC codewords are interleaved to improve burst error tolerance, e.g. 802.3bs is based on interleaving 2x200G FEC codewords.
FEC Lanes	The number of logical lanes that the output of the FEC encoder (and after interleaving if implemented) is distributed across, typically on a symbol by symbol basis. If # FEC lanes > 1, then a deskew and reorder function must be included at the front end of the FEC receiver to correctly reconstruct the aggregate FEC signal before decoding.
AUI Lanes	The number of physical electrical lanes the logical FEC lanes are carried over. # of FEC lanes ≥ # of AUI lanes. Muxing (symbol or bit) can be used to convert between AUI interfaces of different lane widths.
PMD Lanes	The number of lanes at the MDI. The number of AUI lanes at the out of the final stage of muxing must equal the number of PMD lanes.

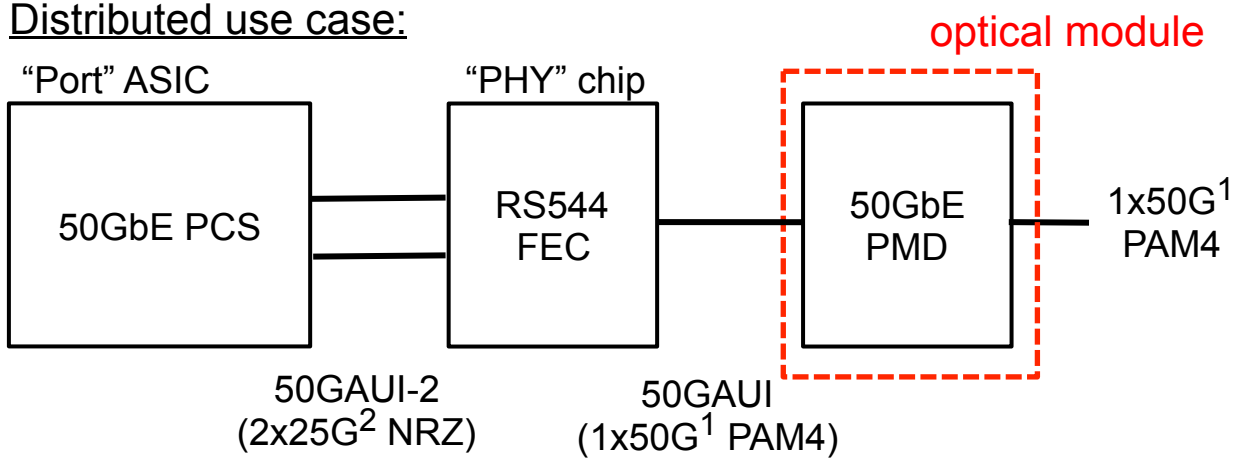
Note: PMA blocks not shown for clarity.

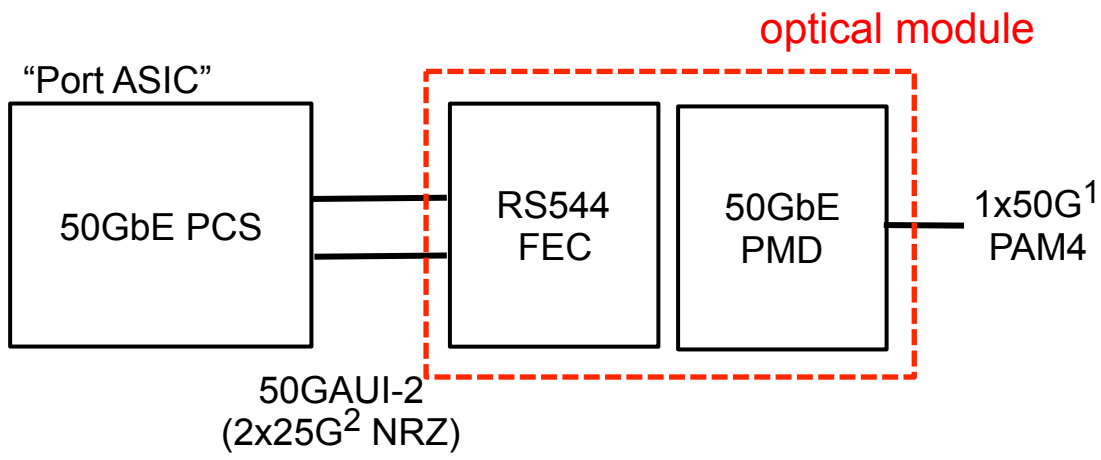
Recap - 50GbE Use Cases (nicholl_3cd_01_0516)

Integrated use case (long term, single lane optimized):



Distributed use case:

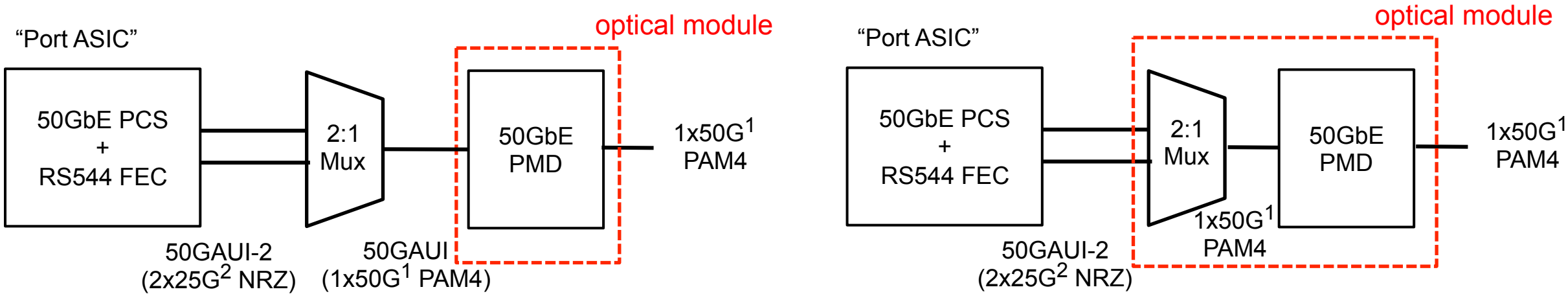




1 = 53.125 Gb/s, 2 = 25.78125Gb/s

Note: PMA blocks not shown for clarity.

50GbE Use Cases (not supported by nicholl_3cd_01_0516)

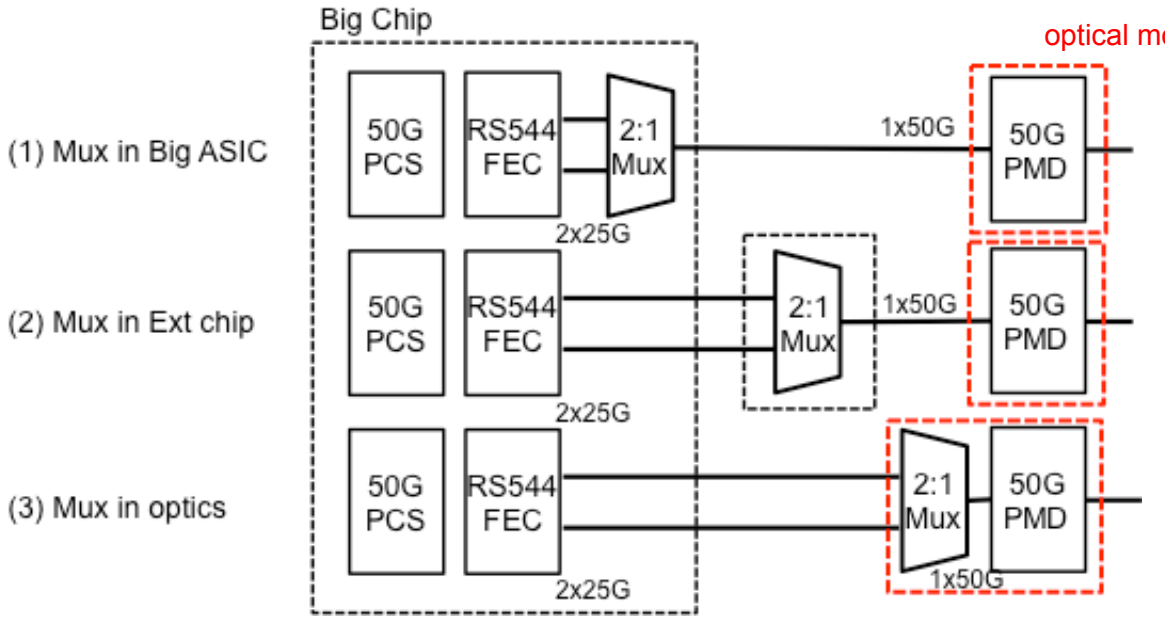


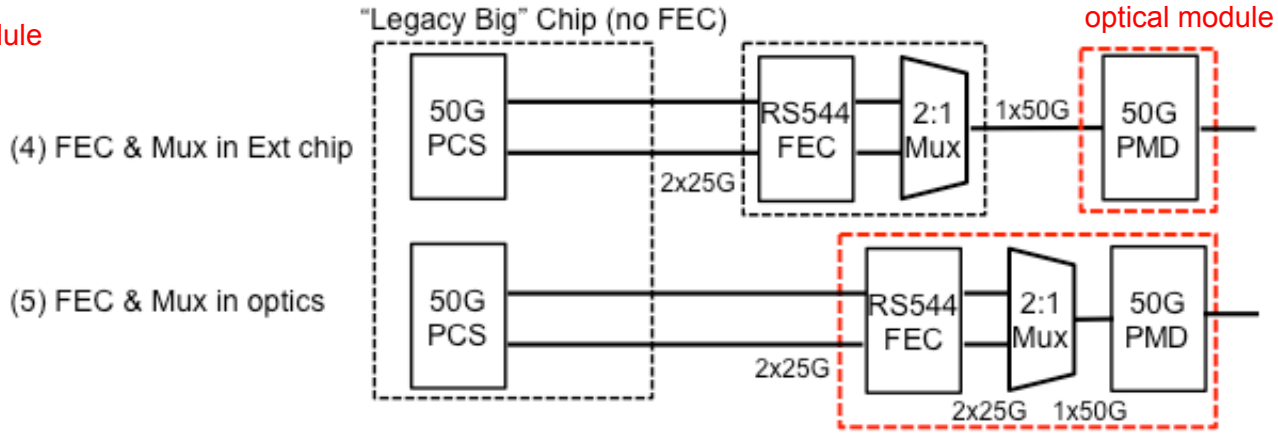
- A FEC distributed to a single 50G FEC lane (nicholl_3cd_01_0516) cannot support these use cases (as it cannot be carried over a 2x25G AUI)
- To support these use cases the FEC must be distributed over 2x25G FEC lanes, with downstream muxing:
  - symbol mux = increase in complexity
  - bit mux = increase in complexity (not as much in optics) + loss in FEC performance

1 = 53.125 Gb/s, 2 = 26.5625Gb/s

Note: All PMA blocks not shown for clarity.

50GbE Use Cases (2x25G Lane FEC Architecture)



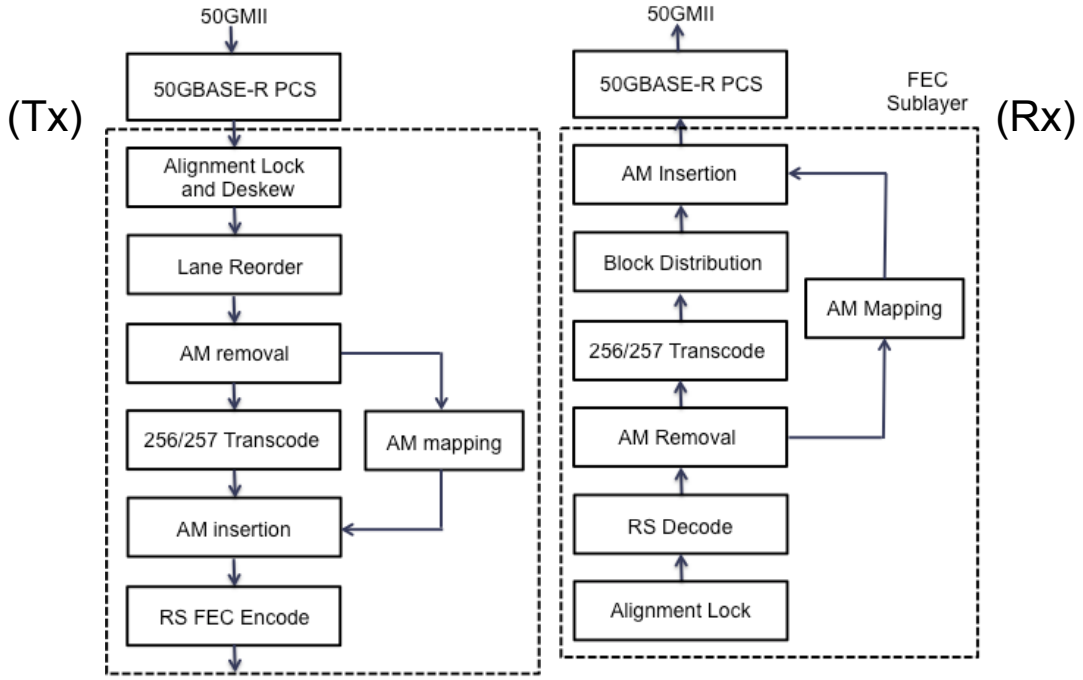


Note: 2:1 Mux could be symbol or bit.

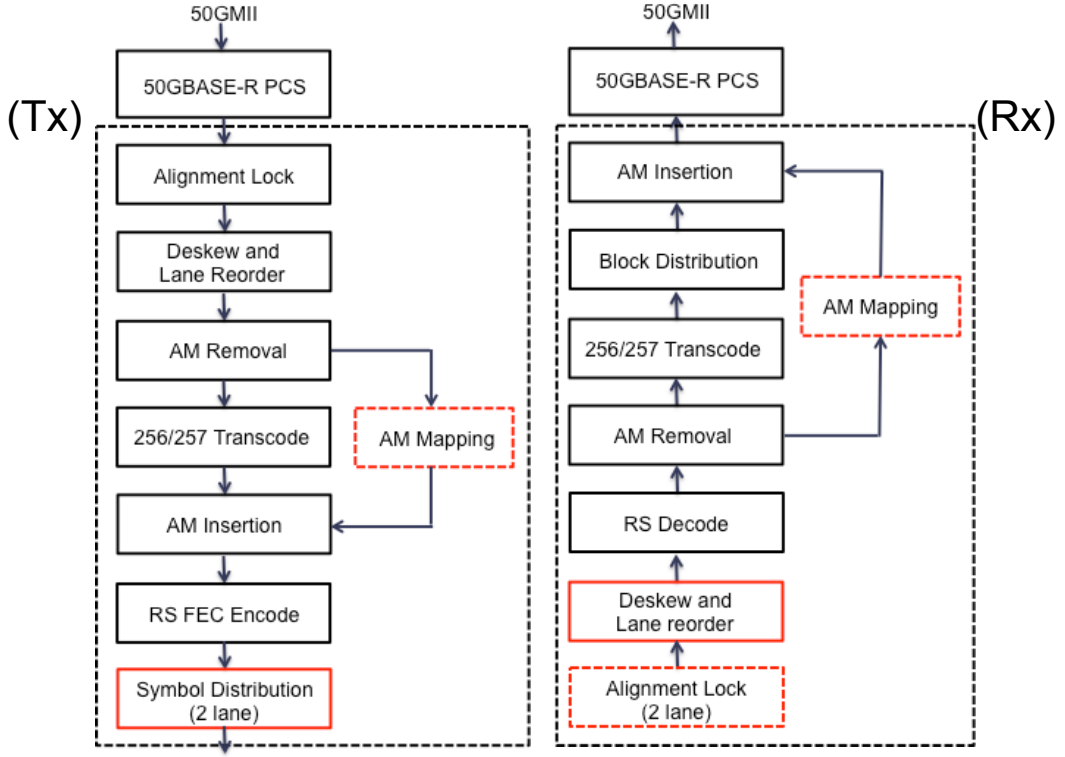
- (1) is the long term volume application, (2)-(5) are all transitory
- (1) contains distribution/deskew in the FEC and a 2:1 mux that is not technically necessary
- Use cases (2) and (3) are enabled by this 2x25G lane FEC architecture
- However it doesn't eliminate the need for an External FEC chip or FEC in optics - use cases (4) and (5)
  - in this regard it is no different to nicholl_3cd_01_0516

Changes to PCS/FEC (1x50G versus 2x25G FEC Lanes)

nicholl_3cd_01_0516 (1x50G FEC Lane)



2x25G FEC Lane Architecture



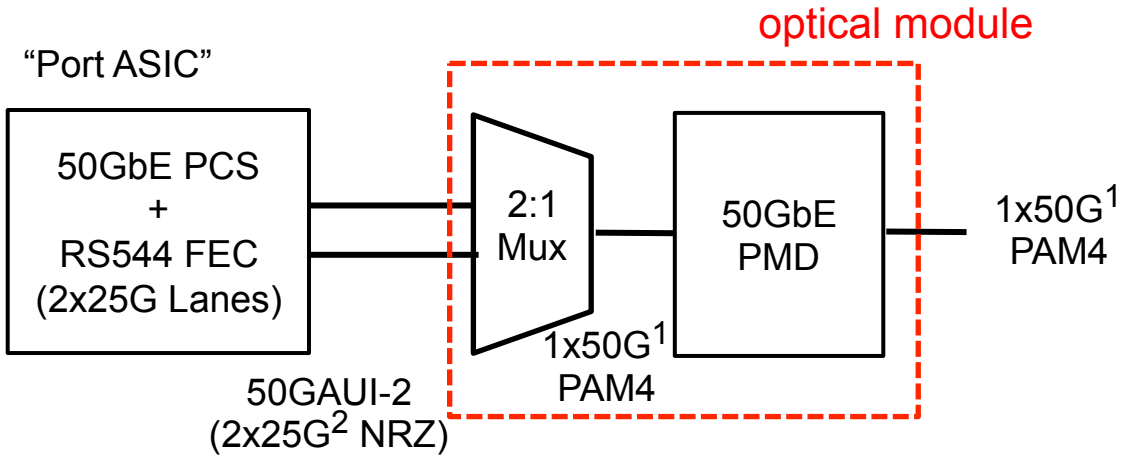
- No change to PCS. Minor changes to FEC sub-layer
- Changes are independent of symbol versus bit muxing
- No impact to latency
- Note: Changes to FEC sub-layer may be even less for 100GbE

Added functionality

Changed functionality

Note: All PMA blocks not shown for clarity.

Symbol versus Bit Muxing ?



Symbol Muxing:

- more complex (codeword alignment)
- solution must be protocol aware
- historically strong pushback
- no FEC performance hit (same as nicholl_3cd_01_0516)

Bit Muxing:

- simple (blind bit muxing)
- solution is protocol agnostic
- FEC performance hit in presence of burst errors (need to quantify)

1 = 53.125 Gb/s, 2 = 26.5625Gb/s

Summary

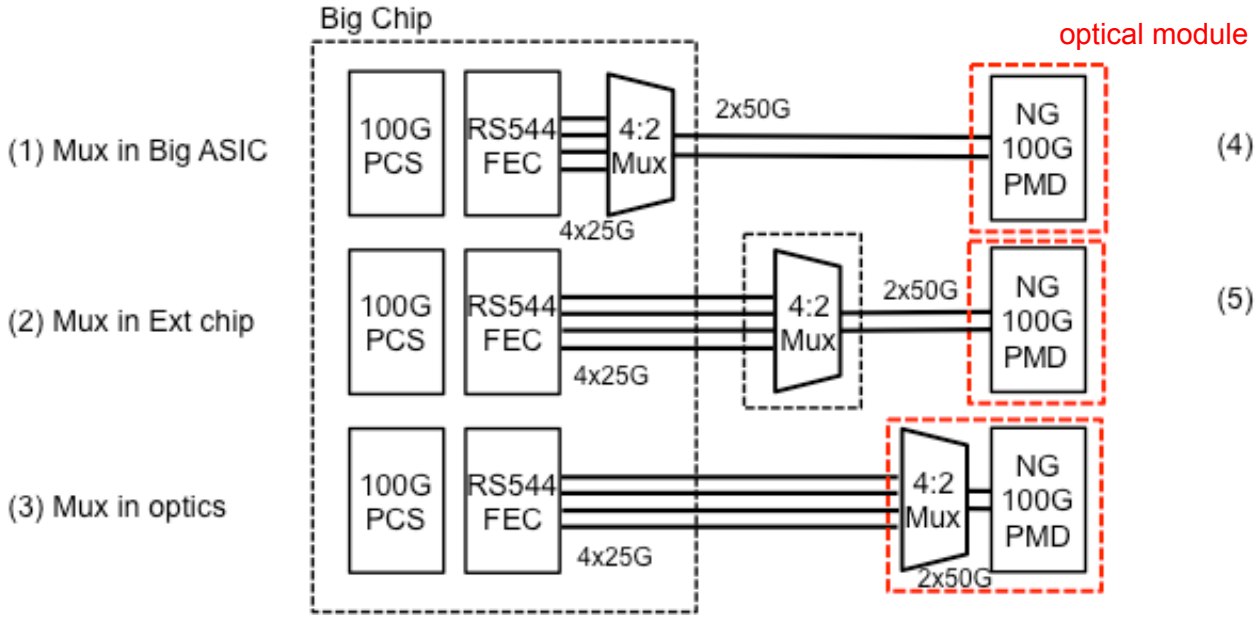
- There is value in a FEC architecture based on 25Gb/s FEC lanes, in that it supports some additional (transitory?) use cases
- The changes to support such an architecture are incremental to the baseline proposed in nicholl_3cd_01_0516
- Need a better understanding of the trade-off between the two different muxing approaches:
  - Impact of implementing symbol muxing in the optical module
  - FEC performance hit in the presence of burst errors for bit muxing
  - Determination of probability of error burst continuing factor “a”

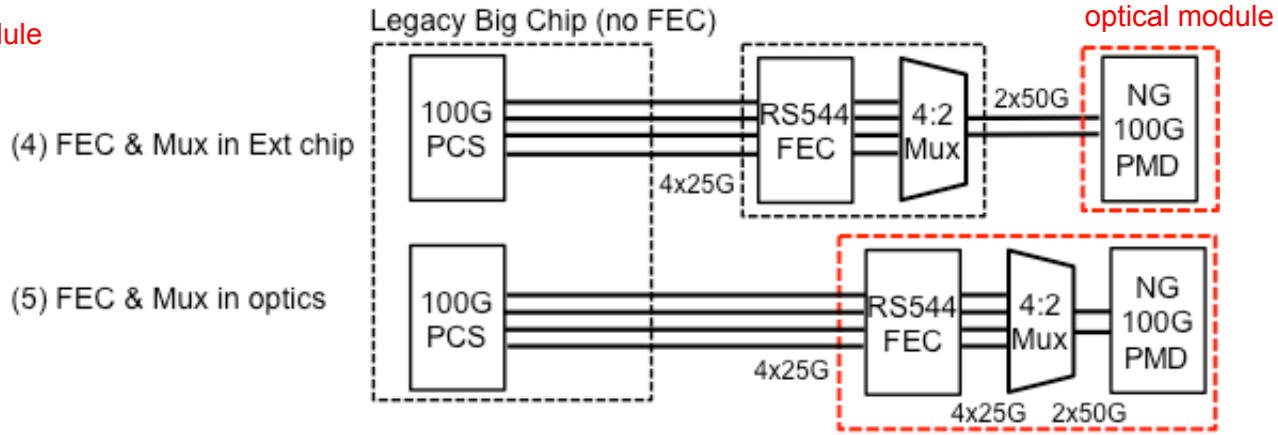
Thank You

Backup

Note: All PMA blocks not shown for clarity.

100GbE Use Cases (4x25G Lane FEC Architecture)



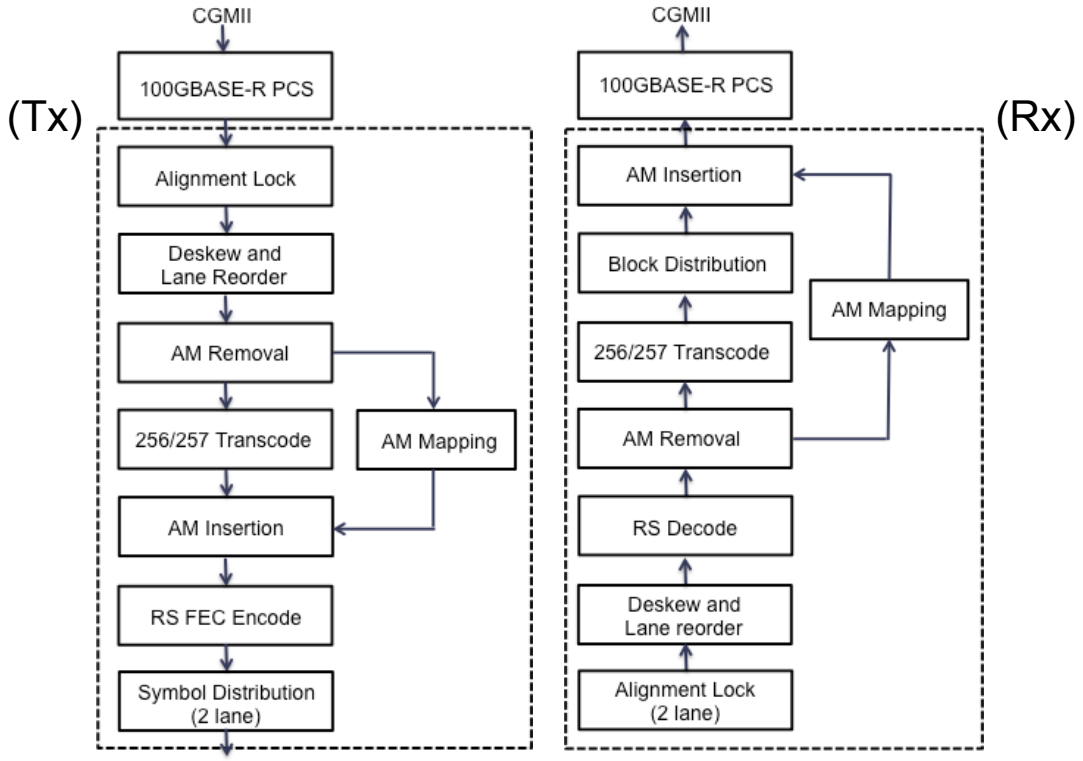


Note: 4:2 Mux could be symbol or bit.

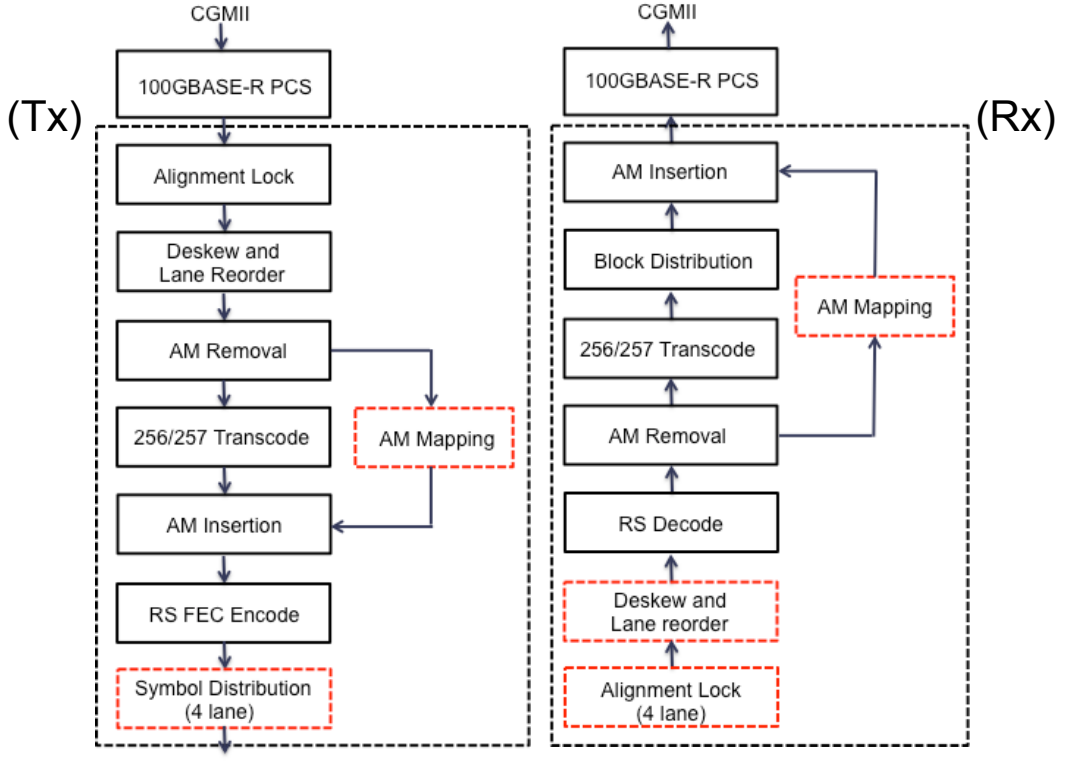
- (1) is the long term, volume application, (2)-(5) are all transitory
- (1) contains distribution/deskew in the FEC and a 4:2 mux that is not technically necessary
- Use cases (2) and (3) are enabled by this 4x25G lane FEC architecture
- However it doesn't eliminate the need for an External FEC chip or FEC in optics - use cases (4) and (5)
  - in this regard it is no different to nicholl_3cd_01_0516

Changes to 100G PCS/FEC (2x50G versus 4x25G FEC Lanes)

nicholl_3cd_01_0516 (2x50G FEC Lane)



4x25G FEC Lane Architecture



- Similar comments to 50GbE
- However in this case 4x25G FEC sub-layer is identical to 802.3bj Clause 91 !

Added functionality

Changed functionality