PCS Baseline proposal for 50GbE and NG 100GbE

Gary Nicholl, Cisco Systems

IEEE 802.3 50GE & NGOATH Study Group ad-hoc, April 27, 2016

Introduction

- Update on my previous presentation:

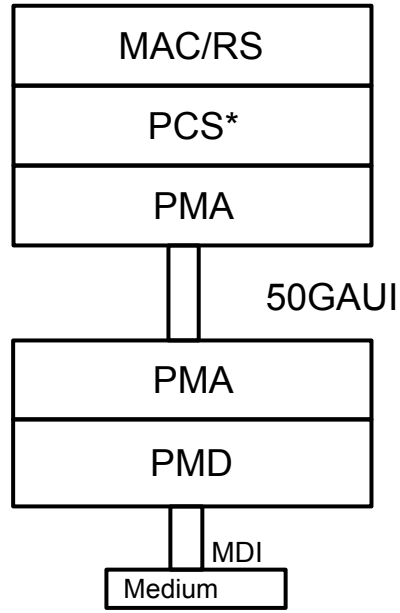
  http://www.ieee802.org/3/50G/public/adhoc/archive/nicholl_041316_50GE_NGOATH_adhoc-v2.pdf

- Presents an alternative 50GbE PCS approach based on a 4 lane MLD architecture (overclocked 40GbE).
- Added some use cases for both 50GbE and 100GbE proposals

50GbE

Recap - Single lane 50GbE PCS Overview

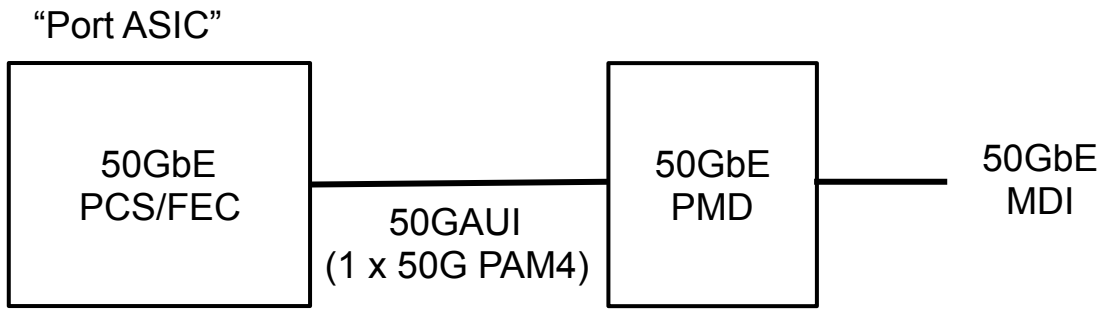
- Single lane PCS
  - optimized for 50Gb/s I/O
  - support for single lane 50Gb/s AUI (only)
- FEC is part of the PCS (no separate FEC sublayer)
  - similar to 802.3bs architecture
- RS (544,514) FEC
- End to end FEC is assumed
  - single FEC instance to cover AUI(s) + PMD
- No FEC codeword interleaving (latency concerns)
- Periodic Codeword Marker (CM) to facilitate FEC codeword alignment



*FEC is part of the PCS sublayer

Note: PMA blocks not shown for clarity.

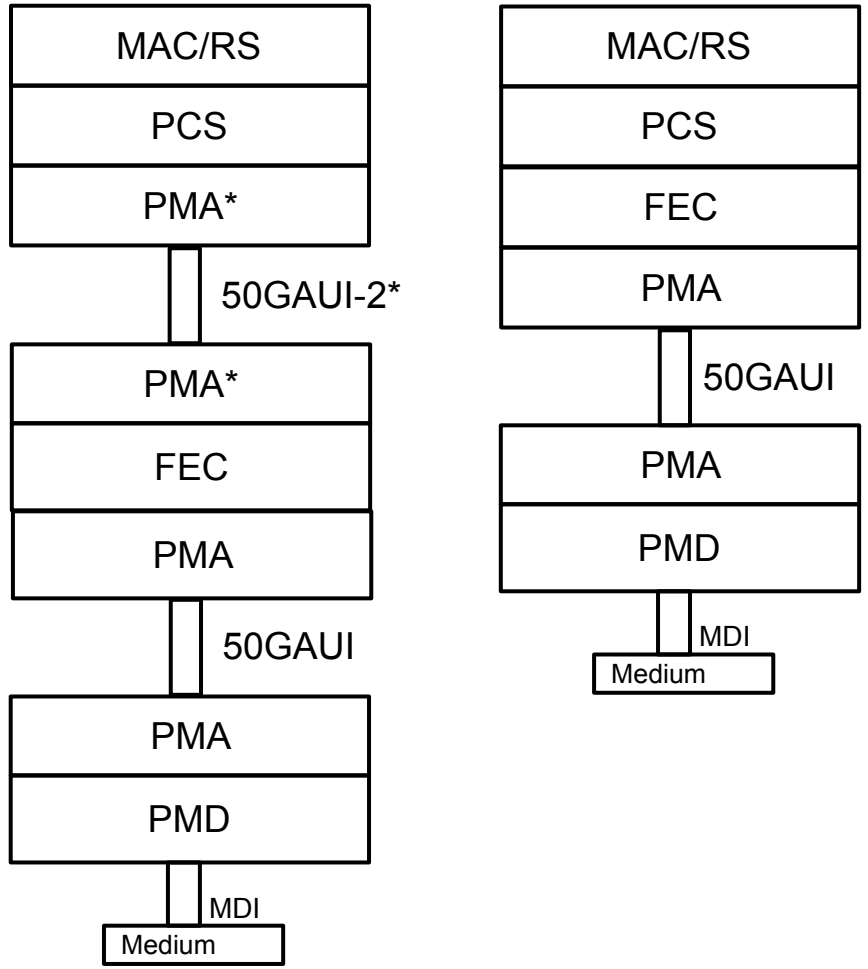
Single Lane 50GbE PCS Use Cases



- port ASIC must have 50Gb/s I/O

Multilane 50GbE PCS Overview

- 4 lane PCS
  - based on overclocked 40GbE PCS (Clause 81)
  - AM spacing modified to support FEC sublayer
  - architecture supports optional AUI-2 /w no-FEC
- Separate FEC sublayer
  - similar to 802.3bj architecture
- RS (544,514) FEC
  - based on 802.3bj (CL 91) but with single FEC lane
  - optimized for 50Gb/s AUI and PMD lane rates
  - no FEC codeword interleaving (latency concerns)
- Enables PCS and FEC to be physically separated (if desired)

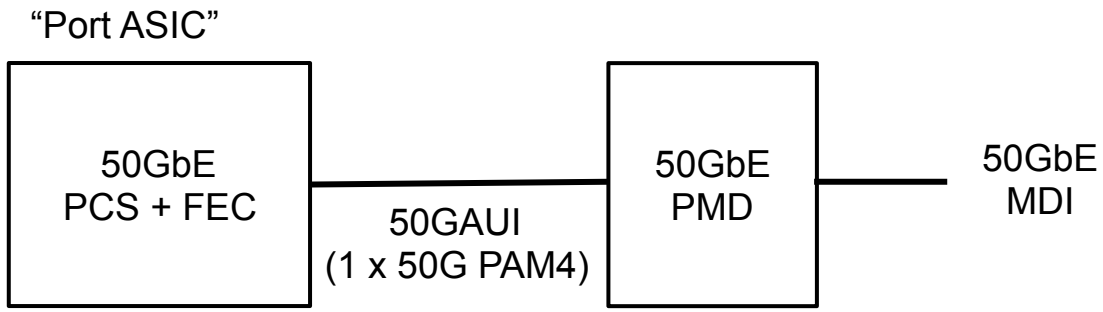


*Optional 50GAUI-2

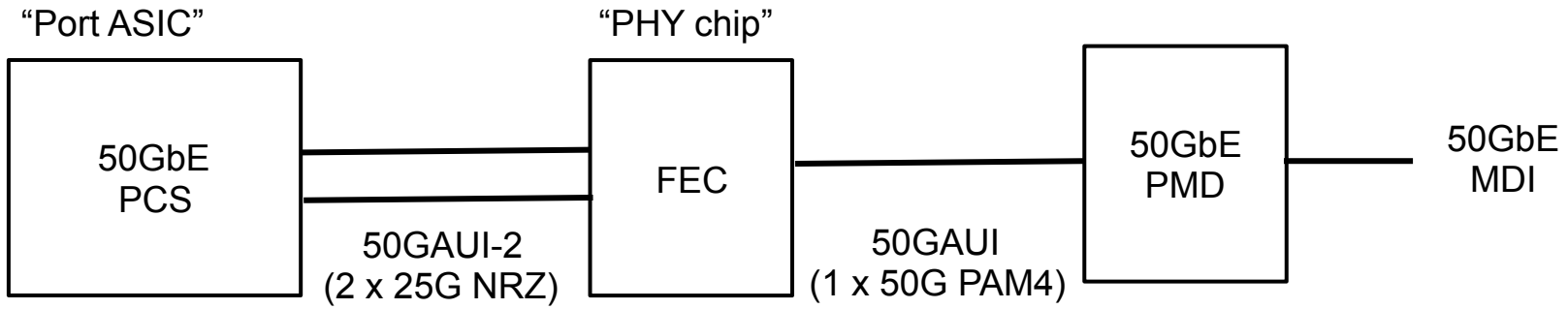
Note: PMA blocks not shown for clarity.

Multilane 50GbE PCS Use Cases

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



Distributed use case:



- port ASIC can start with 25Gb/s I/O and migrate to 50Gb/s over time

Multilane 50GbE PCS Summary

Pros:

- still supports an optimized single lane architecture (with PCS & FEC in port ASIC)
- supports both 50GAUI (1x50G) and optional 50GAUI-2 (2x25G NRZ) interfaces
- enables easy transition from 25Gb/s to 50Gb/s port ASIC IO
- supports ‘bump in the wire’ applications for server ports

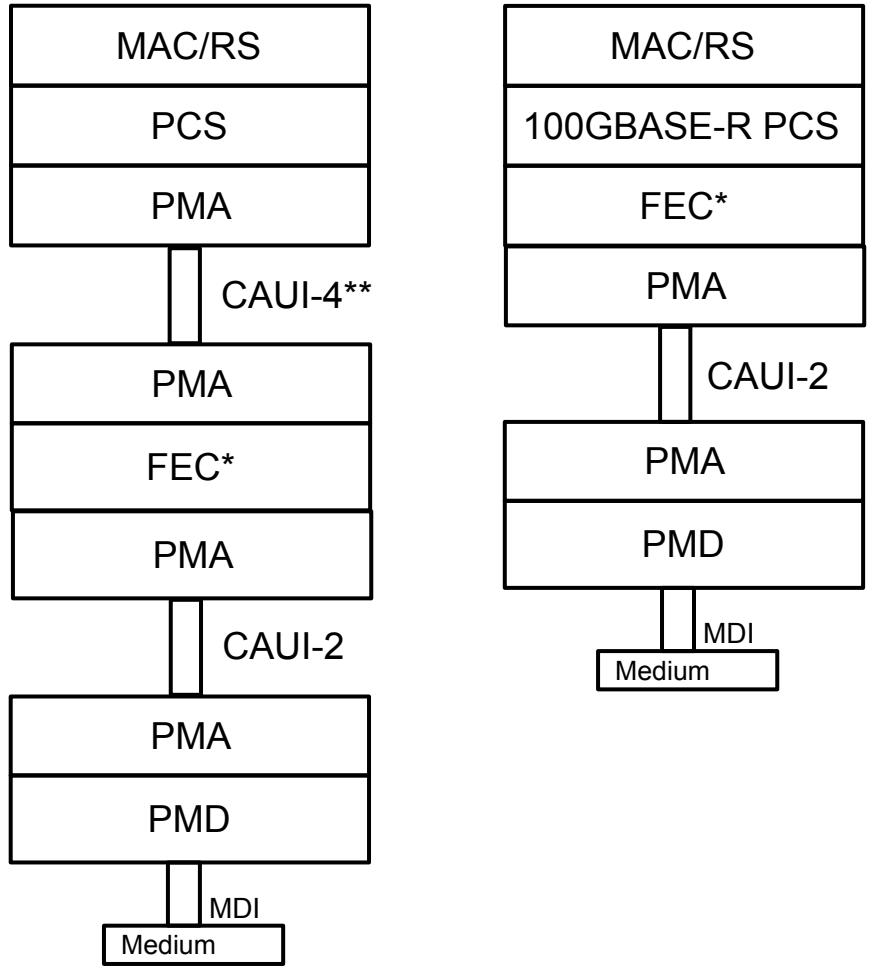
Cons:

- long term the 4 lane MLD functionality in the PCS is redundant

100GbE

Recap - NG 100GbE PCS Overview

- Existing 100GbE (CL82) PCS
  - no changes required
  - supports optional CAUI-4 /w no-FEC
- Separate FEC Sub-layer
  - similar to 802.3bj architecture
- RS (544,514) FEC
  - based on 802.3bj (CL 91) but distributed over 2 FEC lanes
  - optimized for 50Gb/s AUI and PMD lane rates
  - no FEC codeword interleaving (latency concerns)



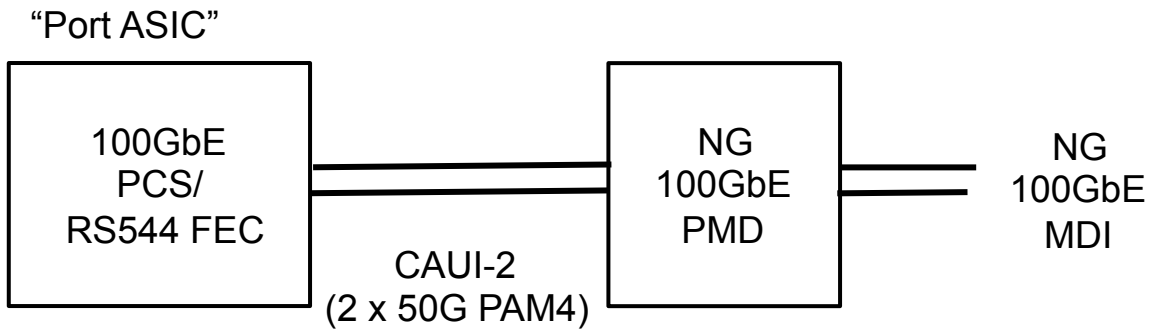
*FEC is a separate sublayer

**optional CAUI-4

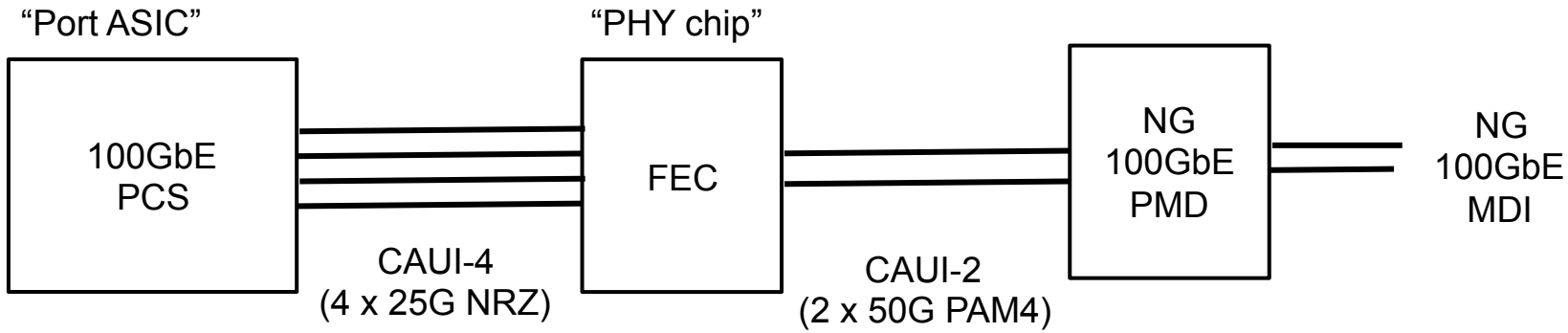
Note: PMA blocks not shown for clarity.

NG 100GbE PCS Use Cases

Integrated use case (long term, 2x50G lane optimized):



Distributed use case:



- can bolt new NG FEC to existing port silicon (either on line card or in module)

Recap - NG 100GbE Backwards Compatibility

- Proposal supports backwards compatibility with legacy hosts:
  - use downspeed serdes (run in 4x25G NRZ mode)
  - reduced bandwidth on new line card (but no different to 1G/10G and 40G/100G transition)
  - requires absolutely no new standards and/or product development
- Proposal supports backwards compatibility with legacy hosts at full bandwidth:
  - new module development with RS544 FEC sublayer installed in legacy host (Rob's Brown Field B)
- Proposal supports backwards compatibility with legacy silicon:
  - new line card with legacy silicon + new (4:2) PHY chip with RS544 FEC
  - this is identical to how RS528 FEC was introduced in transition from 802.3ba to 802.3bj/bm

Summary

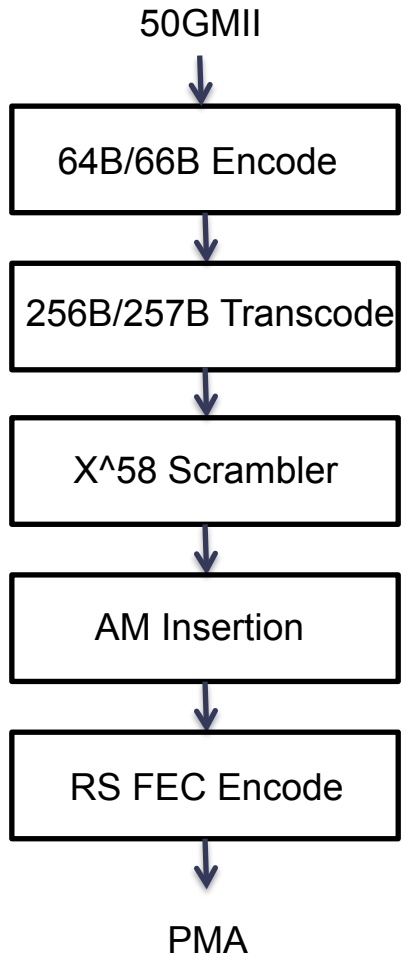
- Baseline proposals presented both 50 GbE and 100 GbE PCS
- Optimized for 50 Gb/s AUI and PMD lane rates
  - new RS544 FEC optimized to run over 50Gb/s lanes only (no muxing from lower rates)
- 100GbE proposal supports optional CAUI-4 Interface (/w no FEC)
- New 50GbE proposal also supports optional 50GAUI-2 interface (w/ no FEC)
- 100 GbE proposal inherently supports a level of backwards compatibility with existing 100 GbE systems
- Any extension to 100 GbE backwards compatibility would be an additional proposal that is incremental to this proposal

Thanks !

Backup

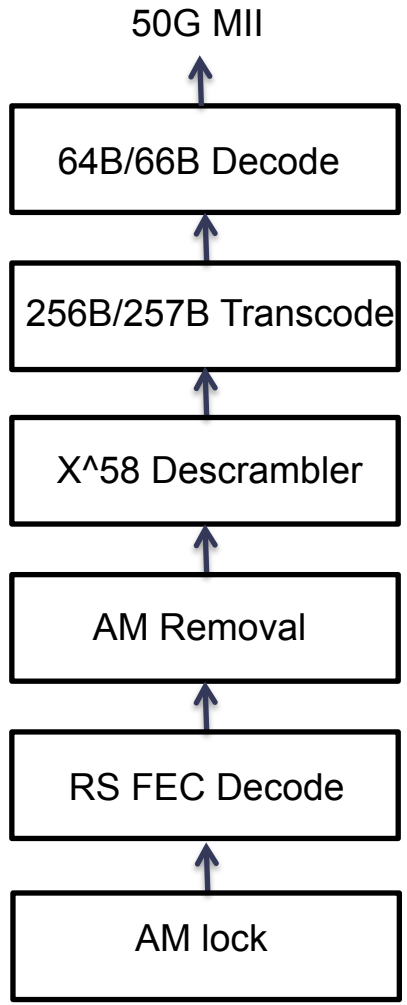
Recap - Single Lane 50GbE Tx PCS Data Flow

- 64B/66B encoder based on Clause 82
- Transcode to 256B/257B based on Clause 91
- Scrambler moved to after the transcoding to simplify the signal flow, standard X^58 scrambler
- Periodic single 257-bit CM Insertion
  - Format and spacing TBD
  - based on CL108 (25GbE)
- RS(544,514) FEC
  - FEC processing as in clause 91
- Support for optional EEE deep sleep
  - based on CL 108
- Supports single physical lane only



Recap - Single lane 50GbE Rx PCS Data Flow

- Reverse of Tx

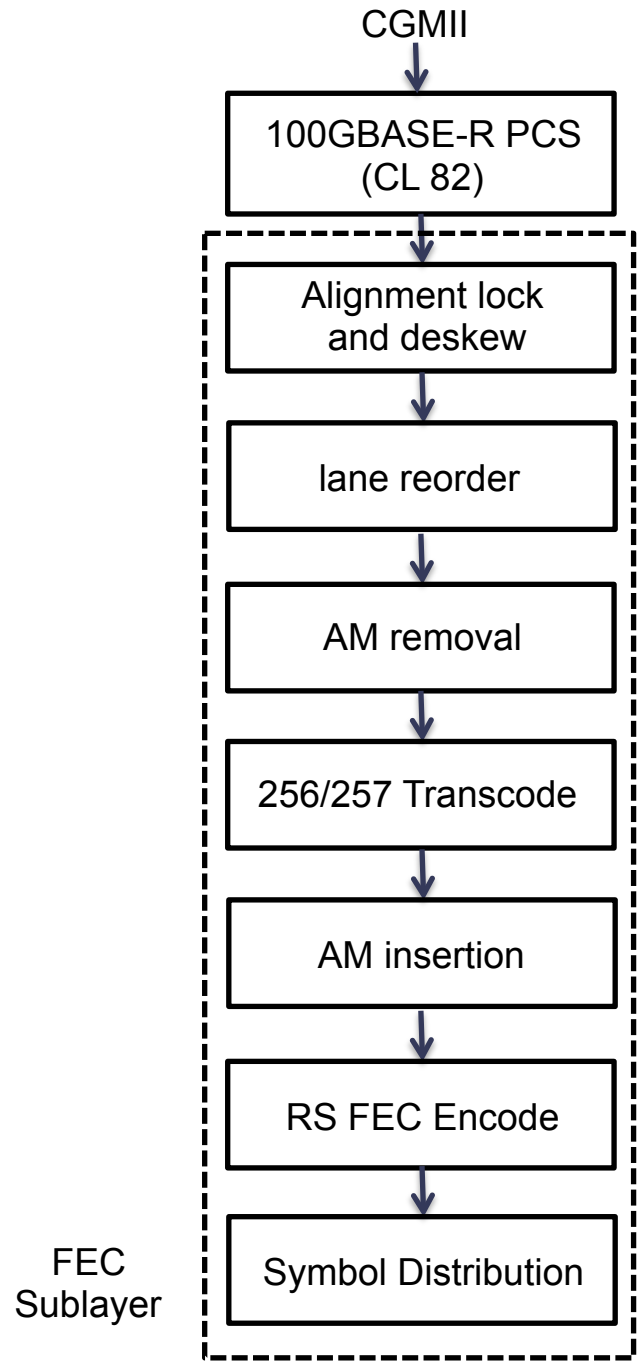


50GbE Open Questions / Things To do

- Need to partition the FEC gain across the different electrical and optical interfaces, and determine the target specs for both.
  - similar to the analysis Pete performed in 802.3bs
  - similar to the analysis Tongtong started in “wang_50GE_NGOATH_01_0316”
  - initial analysis indicates that the 50GAUI and 100GAUI-2 electrical specs may have to be different (tighter) than the current CDAUI-8 specifications in 802.3bs
- Are lower gain / lower latency FEC options (such as RS-528) needed/desired
  - is an end-to-end RS528 FEC a technically viable solution to address any of the objectives ?
  - would mean additional PCS clauses (as the FEC is part of the PCS)
- Is there a valid application for a No-FEC option
  - again this would mean additional PCS clauses (or at least options)

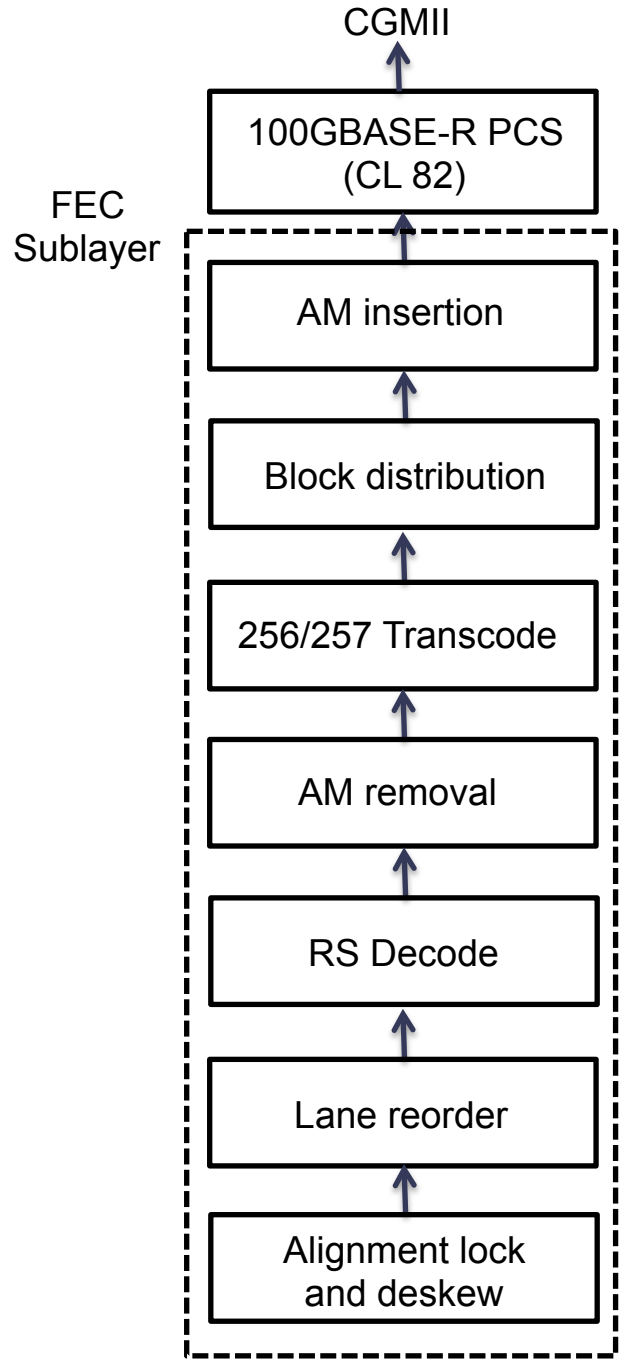
NG 100GbE Tx PCS Data Flow

- PCS data flow same as 802.3ba Clause 81
  - no changes required
- FEC sublayer data flow same as 802.3bj Clause 91
  - FEC symbols distributed over 2 rather than 4 lanes



NG 100GbE Rx PCS Data Flow

- Reverse of Tx



Considerations NG 100GbE Backwards Compatibility

- Proposal supports backwards compatibility with legacy hosts:
  - use downspeed serdes (run in 4x25G NRZ mode)
  - reduced bandwidth on new line card (but no different to 1G/10G and 40G/100G transition)
  - requires absolutely no new standards and/or product development
- Proposal supports backwards compatibility with legacy hosts at full bandwidth:
  - new module development with RS544 FEC sublayer installed in legacy host (Rob's Brown Field B)
- Proposal supports backwards compatibility with legacy silicon:
  - new line card with legacy silicon + new (4:2) PHY chip with RS544 FEC
  - this is identical to how RS528 FEC was introduced in transition from 802.3ba to 802.3bj/bm

NG 100GbE Open Questions / Things To do

- Need to partition the RS544 FEC gain across the different electrical and optical interfaces, and determine the target specs for both.
  - similar to the analysis Pete performed in 802.3bs
  - similar to the analysis Tongtong started in “wang_50GE_NGOATH_01_0316”
- RS528 FEC is still likely to be supported in new silicon, even with 50G I/O
  - used when running in downspeed CAUI-4 mode
  - is it also possible to run RS528 FEC across a 2 lane 100GbE link, and if so what are the performance implications
  - would it support any of the current objectives ? If not is there any interest in adding new objectives that could be supported (but essentially means an additional set of both AUI and PMD specifications).