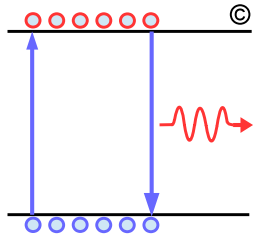


PCS Consideration for 50 GbE and NG 100 GBE

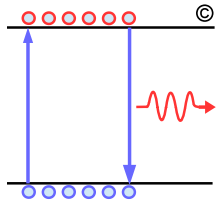
Ali Ghiasi

Ghiasi Quantum LLC

NGOATH AdHoc Meeting

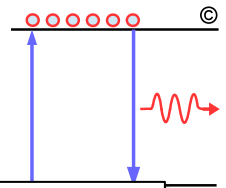
Feb 17, 2016

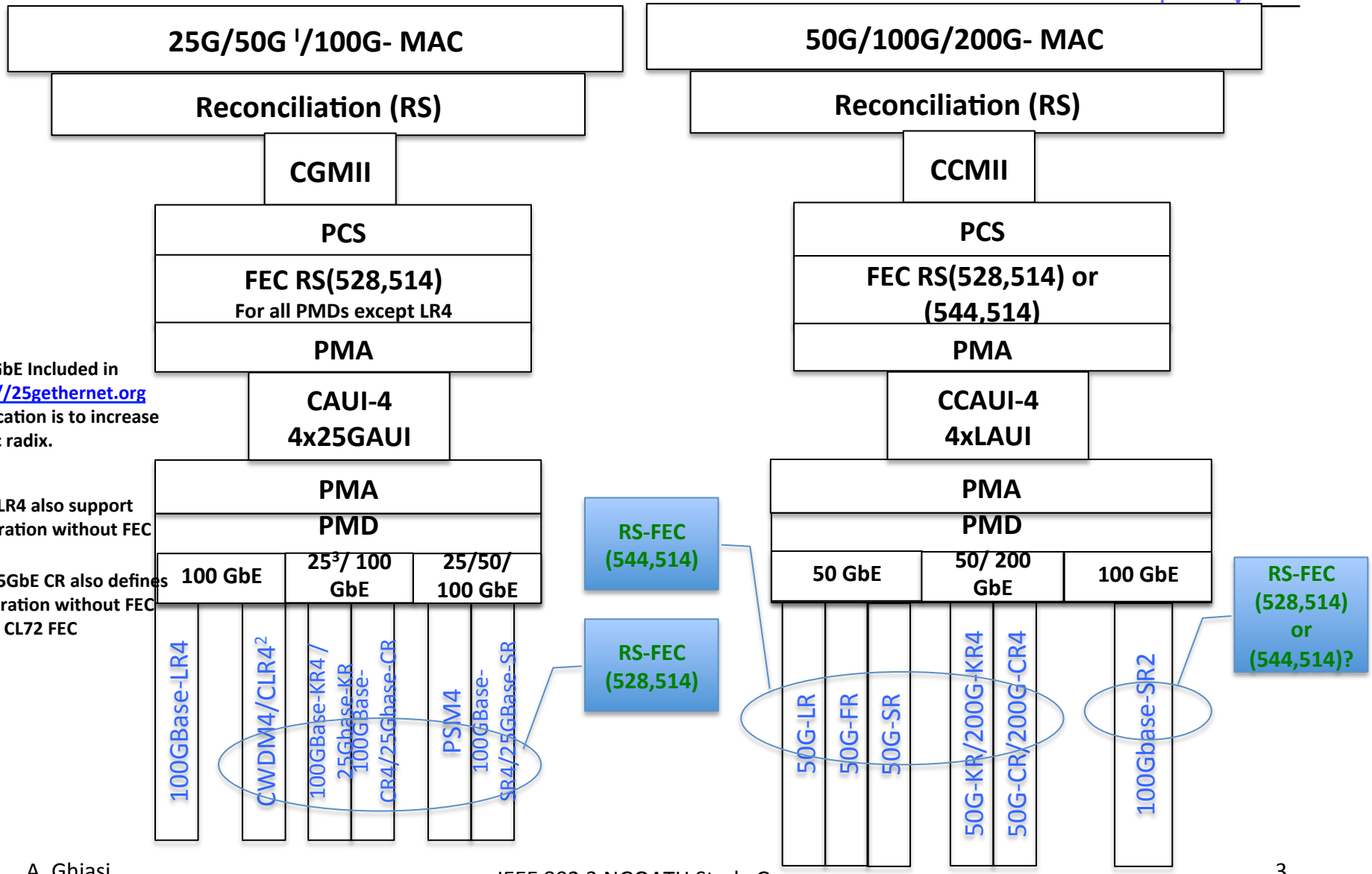
Observation on 50GbE, 200 GbE, and NG 100GbE PMDs



- **50 GbE and 200 GbE are complimentary set of standards just as we observed in the market place the complimentary nature of 25GbE/100 GbE**
  - Current generation of switch ASIC offer 4x25 GbE breakout for small incremental cost
  - Next generation switch ASIC will offer 4x50GbE breakout for same economics
  - 200 GbE assumes 25G PCS so there is no reason for 50 GbE not to use 25G PCS lane
- **NG 100 GbE PMDs attributes and requirements**
  - Currently with the increase in volume the market is enjoying significant cost reduction for 100 GbE PMDs such as 100GBase-SR4, PSM4, and CWDM4/CLR4
  - Cost may not be the main driver to define NG 100 GbE PMDs with exception of CAUI-2
  - Currently defined 100 GbE PMDs will require inverse-mux with introduction of 50G ASIC IO
    - A PMA-PMA device could address any I/O mismatch
    - Simplest form of PMA/PMD implementation occurs for the case when # of electrical lanes = # of optical lanes/λ
  - Do we need with every generation of electrical I/O 25G, 50G, 100G introduce new 100 GbE PMDs which are optimized for given generation of ASIC but not backward compatible?
  - The decision to define new optical PMD should not be taken lightly to save a PMA-PMA mux!

Current 25/50/100 GbE PMDs and NGOATH PMDs in the Current Objective



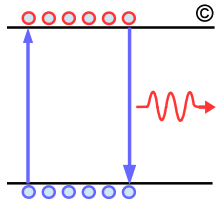


1.50 GbE Included in http://25gethernet.org application is to increase fabric radix.

2. CLR4 also support operation without FEC

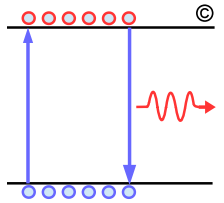
3. 25GbE CR also defines Operation without FEC And CL72 FEC

PCS Synergy Between 50, 100, 200G, and 400 GbE Is Advantages



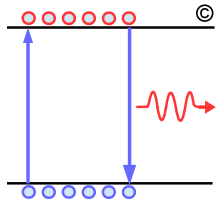
- **400GbE is based on 25G PCS lanes with total of 16 PCS lanes**
  - Can support 16x25G, 8x50G, 4x100G, 2x200G, or 1x400G
- **200GbE uses 25G PCS following 400 GbE but with total of 8 PCS lanes**
  - Can support 8x25G, 4x50G, 2x100G, 1x200G
- **NG 100GbE will be balancing act between backward compatibility and synergy with 50/200 GbE**
  - If we abandon backward compatibility then 4 PCS lanes each at 25G is natural
  - If we want backward compatibility then we need to stay with 20 PCS lanes and KR4 FEC RS(528,514)
- **50 GbE should follow 200GbE with 25G PCS lane since 50GbE will co-existence with 200 GbE as part of break out**
  - Group charter is defining single lane 50G PMD
  - Does not mean optional AUI should not support 2x25G electrical
  - 25G PCS is common and natural for 50 GbE
- **Broad market potential is better satisfied by considering compatibility with 200/400 GbE and 100 GbE!**

Consideration for 50 GbE

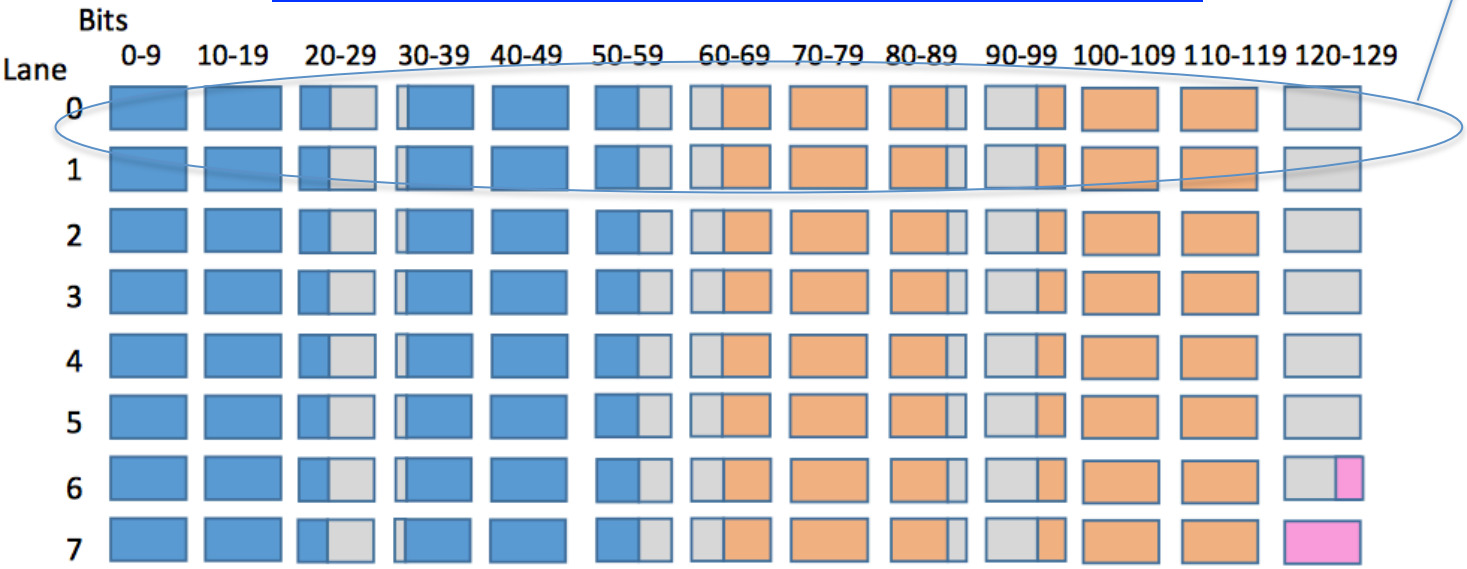


- **Should have a common PCS and FEC to support all PMDs**
  - Frame format and link performance is optimized for single operation
  - At the same time lets not penalize 50G PMDs
    - Lets assume LAUI-2 BER of 1E-12 or better
    - Lets assume symbol mux so there is no penalty for using LAUI-2
  - But lets provide PCS mechanism to support LAUI-2 electrical by enabling early adopters
    - 2x25G – 1x50G LAUI PHY requiring symbol mux is a trivial feature
- **To support backplane and CR links with 30 dB of loss even after tightening the channel parameters even RS-FEC(544,514) only provides marginal COM**
  - Due to latency FEC interleaving likely not an option
  - Likely need all the FEC gain we can get so ruling out bit mux is reasonable
  - Need to consider broad market potential and technical feasibility of low latency link.

50 GbE PCS format Should be Based 200 GbE PCS Format



- **Currently proposed PCS format per**
  - http://www.ieee802.org/3/bs/public/adhoc/logic/feb9_16/gustlin_01_0216_logic.pdf



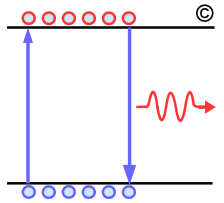
Bases for 50 GbE PCS

- Common marker (2*24)
- Unique markers (2*24)
- Pad (3*64+68 bits)
- Next 257-bit block

- Proposed 200 GbE is base of 4x257b blocks, pad is filled with free running PRBS9
- 50 GbE can be based on 1x257b blocks, pad is filled with free running PRBS9
- 2 PCS lanes can be formed with 2 AM as following:

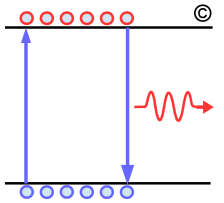
+ Pad bits

Consideration for Next Gen 100 GbE PMD

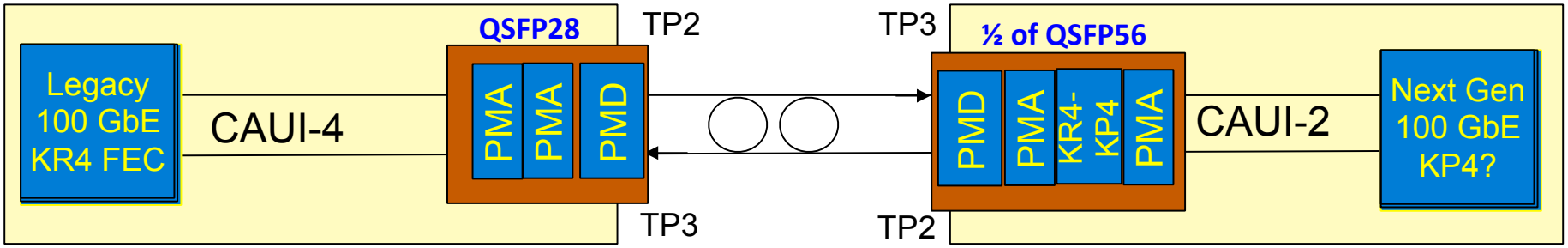


- **With in the NGOATH following PMDs are within the scope**
  - 100Gbase-SR2
  - Optional CAUI-2
- **Main decision for the group is**
  - Make 100 GbE NGOATH backward compatible by using 5G PCS lanes and RS-FEC (528,514)
    - Fully compatible with current 100 GbE PMDs 100GBase-SR4, CAUI-4, 100G-PSM4, and 100G-CWDM4/CLR4
    - Above PMDs are just moving into volume deployment in 2016 and system build over the next two years will only have CAUI-4 and RS-FEC(528,514)
    - If additional FEC gain is needed symbol muxing is reasonable to assume
    - If the NGOATH PMDs can be satisfied with (528,514) FEC then the choice of PCS is obvious and should be 5G PCS lane
  - If we can not satisfy NGOATH PMDs with RS-FEC(528,514) then we might as well leverage synergy with 50/200/400 GbE PCS
    - PCS based on 25G lanes
    - FEC based on RS-FEC(544,514)
    - If we are defining new PCS/FEC then it should support [4,2,1] lanes
  - Defining 100 GbE NGOATH based on 20 PCS lanes with RS-FEC(544,514) is neither backward or forward compatible!

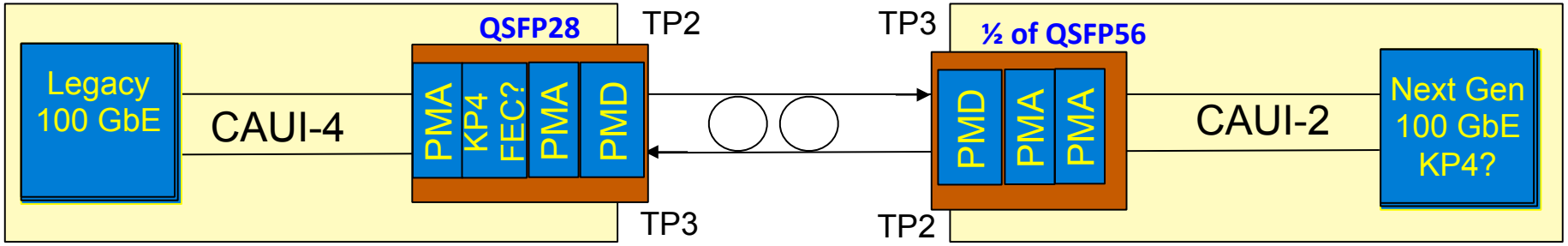
The Challenge with 100 GbE Next Gen PMDs



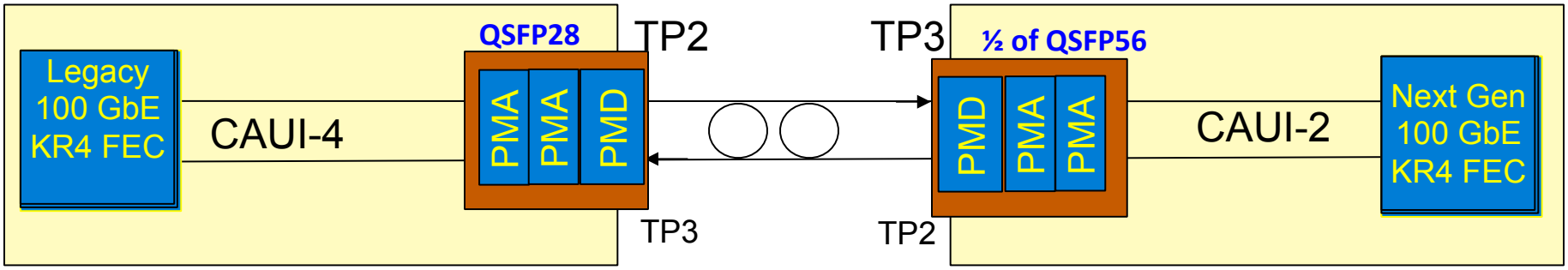
- **Approach support existing 100 GbE PMD**



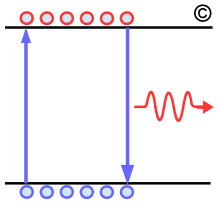
- **Approach to support new 100 GbE PMDs**



- **The simplest approach if feasible is to define new 100 GbE PMDs based on KR4 FEC.**



Summary



- **With proposed 200 GbE PCS having 8 lanes it is natural to have 50 GbE to have 2 PCS lanes**
  - Common ports will support 50/200 GbE through QSFP56 breakout
  - Defining 50 GbE based on 2 PCS the eco-systems gets enabled by current generation FPGA
  - If LAUI-2 supports two lanes then 50 GbE potentially could be supported with PMA-PMA mux in the module
    - NGOATH PMDs and PCS should be optimized for operation based on single lane
    - It is reasonable to assume the LAU-2 PMA-PMA device may need to perform symbol mux, etc
- **Next Gen 100 GbE PMDs can be based on 200 GbE PCS/FEC or it can be defined to be backward compatible using Clause 82 PCS and KR4 FEC**
  - The advantage of using common FEC for 100 GbE and 200 GbE is to achieve identical performance for a PMD operating in full rate or break out mode
  - Considering the investment made in current 100 GbE PMDs backward compatibility should an important consideration if the 100 GbE NGOATH PMD can be satisfied with RS-FEC (528,514)
  - If 100 GbE NGOATH PMDs require RS-FEC(544,514) then lets build synergy with 200/400GbE by assuming 25G PCS lanes
- **The 802.3 need to balance cloud applications driven by fork lift upgrade as well as synergy and compatibility across Ethernet eco-system.**