IEEE 802.3 25G Ethernet SG – Arch Ad Hoc Layering and Gaps

Eric Baden (ericb at broadcom com)
Yong Kim (ybkim at broadcom com), presenting
Cedrik Begin (cbegin at cisco com), presenting

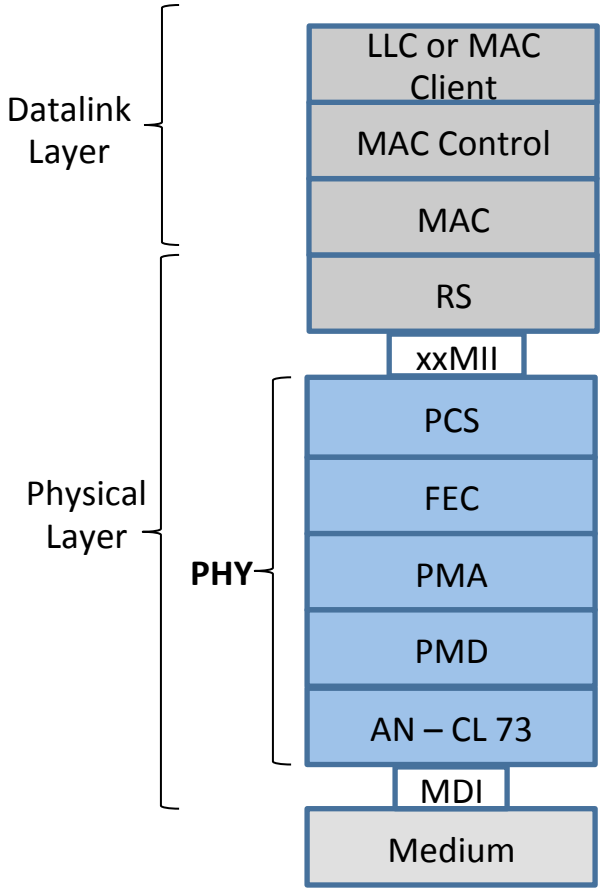
25G PCS Thoughts - recap

- Recap from Sept Interim (not to revisit)
  - Both 3m and 5m reach adopted as objectives (implicit ToR and InterR)
  - FEC/no FEC (implicit sub-set objectives of latency, cost, compatibilities
- Views
  - 10G speed up
  - 100G (.3 bj) quarter lane use
- Desires
  - NICs – implementations for 10G/25G and 40G
  - Switches – implementations for 100G/40G/25G and 10G

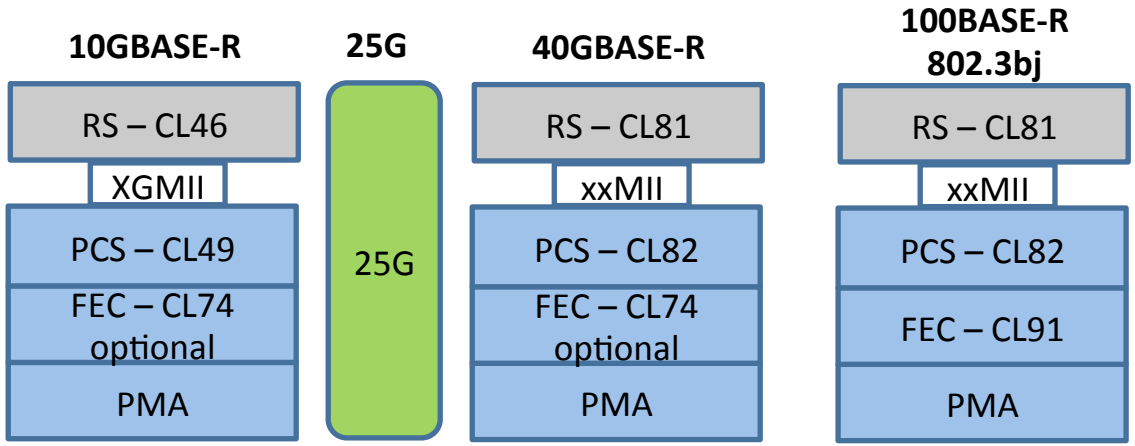
General and Common Ideas - Recap

- 64/66B.
- Lane rate of 25.78125G
- Alignment Marker eases the use of FEC (not FEC capability).
  - BIP has benefits. Bug-fix category or nice to have?
- Optional Auto-negotiation determines use of FEC and training, among other things.

[Sub-]Layering



- 25G RS/PCS/FEC considerations.

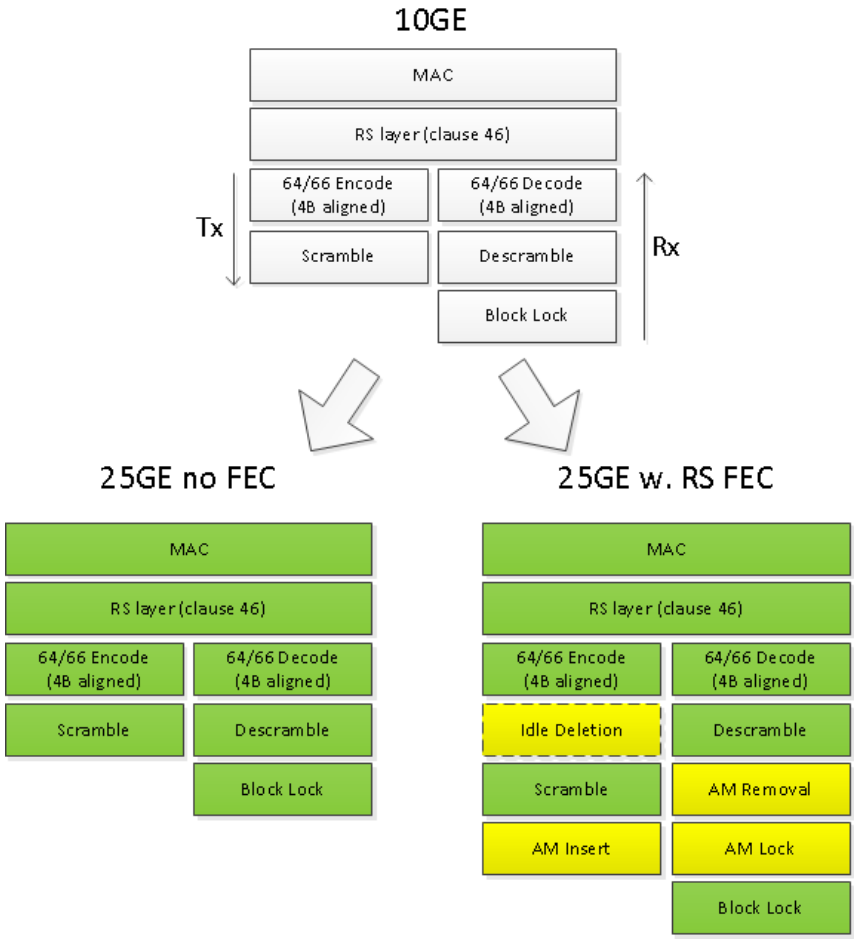


[Sub-]Layer Elements

- Closer look at the data path elements of 10GBASE-R, 40G/100G BASE-R, and recent . 3bj work.
  - Examine RS/PCS/FEC datapath elements adopted for 25G Ethernet use, individual clause basis and also together.
  - Evaluate the choices for relevancy, technical merits, and ease of implementation.

Details of 25G Sub-Sub-Layering considerations

25GE PCS using 10GE (CL49) building blocks

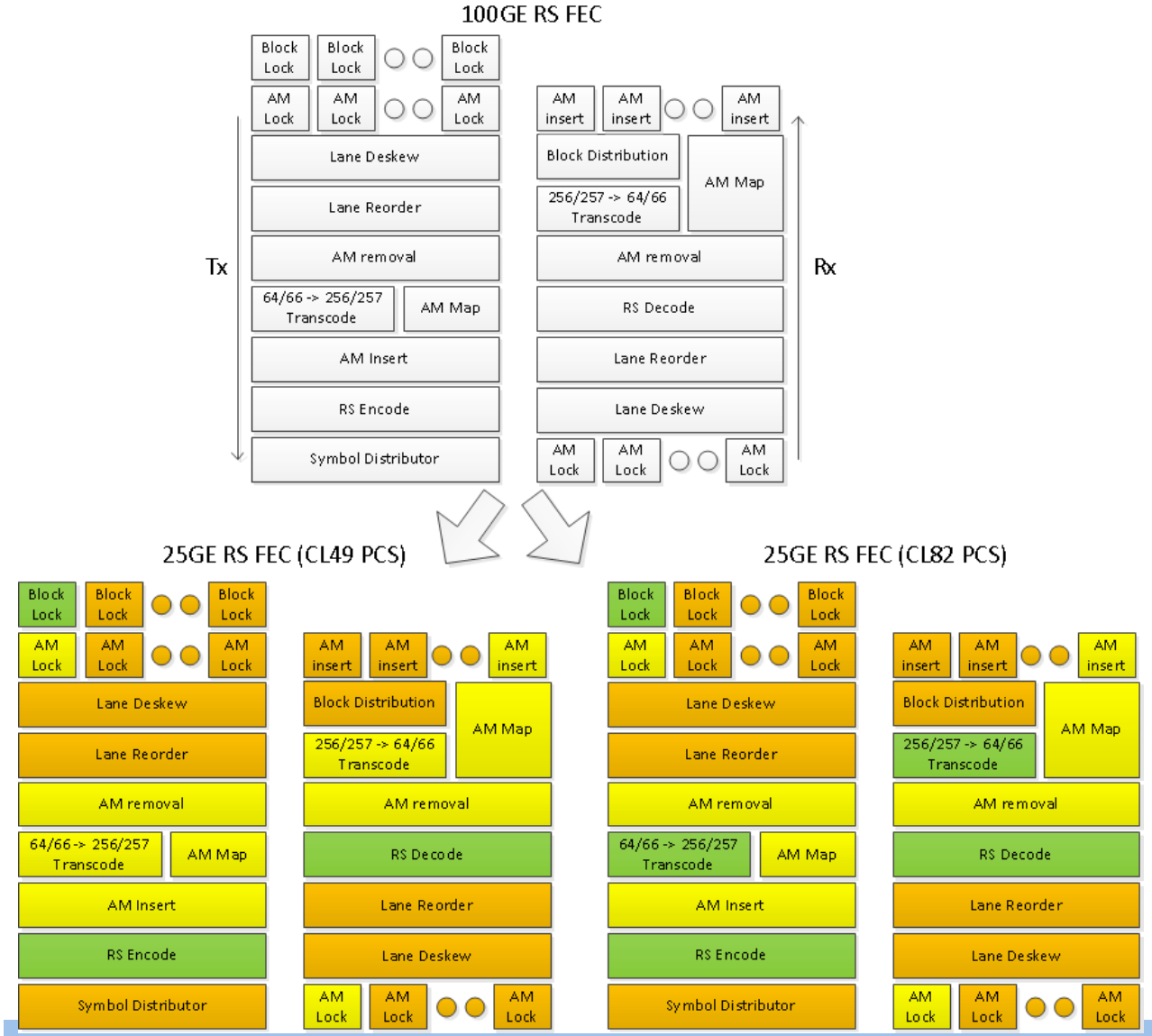


- 4 byte MII (CL46)
- For a 25GE without FEC can use 10GE function as is, i.e. complete reuse (simply run 2.5x faster).
- To aid RS FEC, would add alignment marker insertion and removal in the 25GE PCS. (yellow blocks)

25GE PCS using 40/100GE (CL82) building blocks

- 8 byte MII (CL81).
- Some function reuse, however would remove (orange blocks):
  - multiple per lane logic
  - block distribution and reorder/deskew.
- AM insertion/removal logic would need to change (yellow blocks) in order to reflect different rates of AM insertion/ removal

Changes to RS FEC (CL91) for 25GE (8B vs. 4B)



- For both options would remove (orange):
  - Per lane logic
  - Block distribution and deskew logic.
- For both options would need to change AM related logic to reflect difference in number of AMs and periodicity (yellow).
- Only difference between the two options is that the clause 49 based option would need the transcoders to not restrict the transcoding of its additional block codes.

Summary

- Clause 49 is the better starting point for a 25GE PCS.
  - Even in the case where an alignment marker is inserted to aid the RS FEC
- Changes are required to clause 91 FEC, whether or not the 25GE PCS is based on clause 49 or clause 82
  - Magnitude of changes are equivalent.

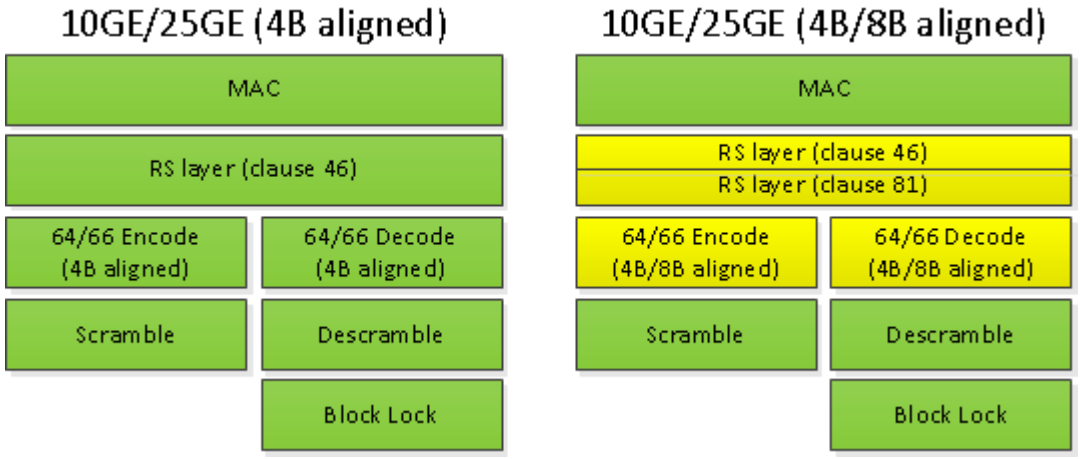
25G directions with optional FEC

RS/PCS/FEC	10G	25G without FEC	25G with FEC	40G	100G
Block Coding			64/66B		
Lanes	1	1	1	4	4
RS	CL46 (4B)	CL46 (4B)	CL46 (4B)	XLGMII (8B)	CGMII (8B)
PCS	CL49	CL49	CL49	CL82	CL82
Align M	-	-	Y	Y	Y
Trans Code	-	-	256/257B	N/A	256/257B
Reach		3+ m	5+ m		
Latency		Low	High		
Optional CL74 FEC Use (TBD)	Y	Y	Y	Y	Y

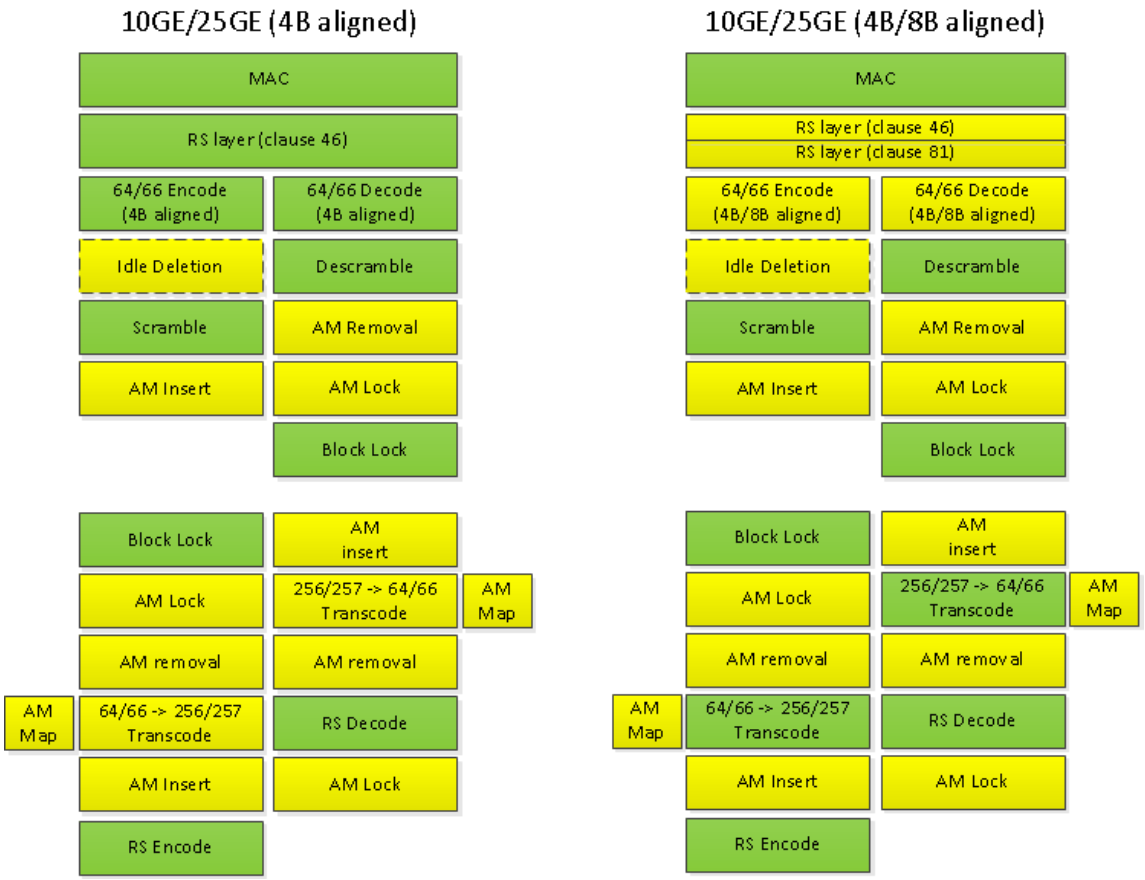
THANK YOU!

Implementation of a 25GE/10GE capable port /wo FEC

- In the case where FEC isn't required, one could build a port capable of 25GE/10GE where the only difference is the frequency if both support the 4B alignment.
- Different byte alignment per port would require:
  - 2 different RS layers.
  - 64/66 encode/decode would need 2 modes depending on the port speed.



Implementation of a 25GE/10GE capable port /w FEC



- In the case where RS-FEC is required, tradeoff would be:
  - Slight change to transcode,

    versus,
  - 2 RS layers + 2 64/66 encode/decode modes.