Background on 25GbE PCS/FEC Baseline Proposal (baden_3by_01b_0115)

Gary Nicholl, Cisco;

Eric Baden, Broadcom

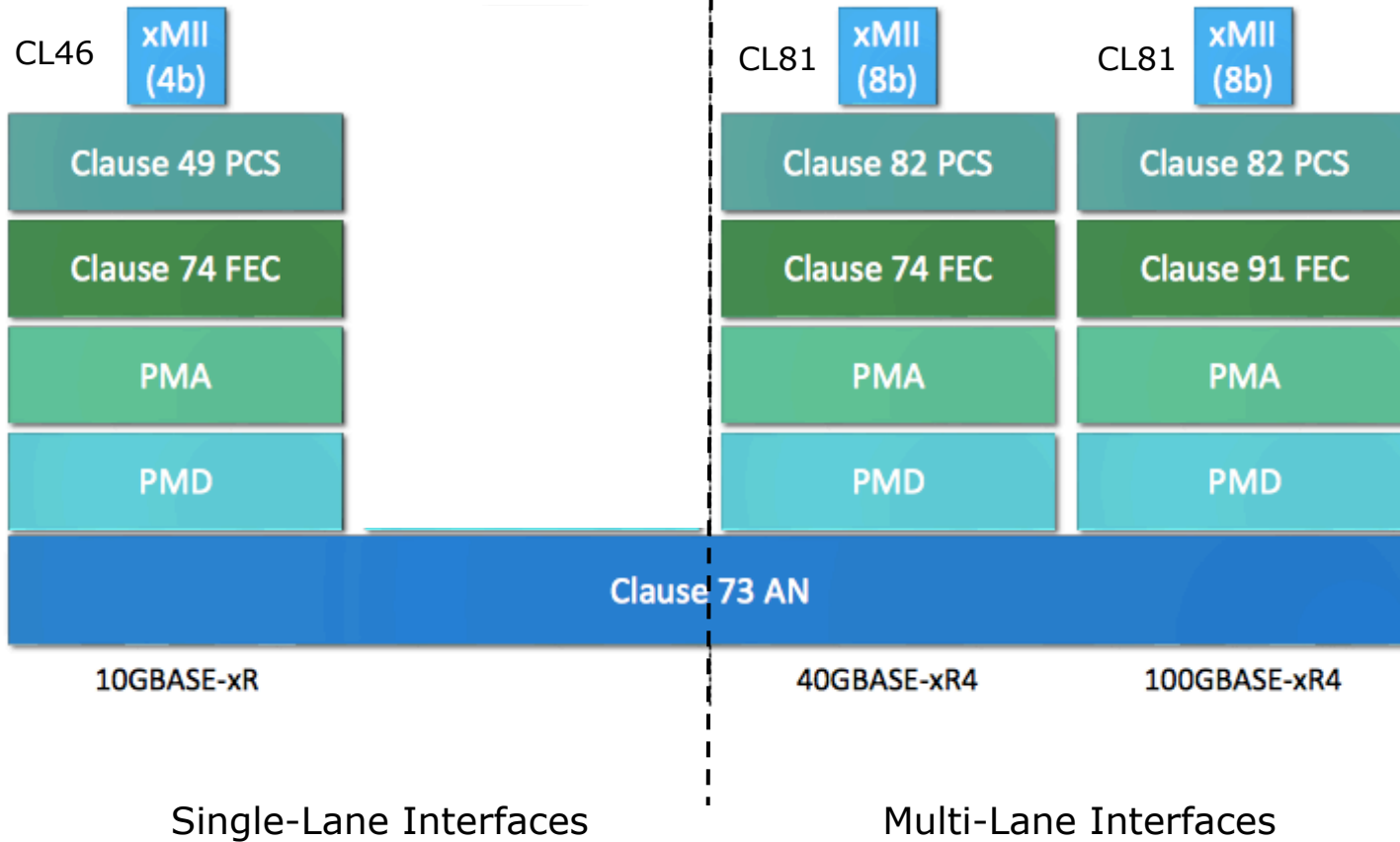IEEE 802.3by TF, Atlanta, Jan 14-16, 2015

Introduction

- The candidate baseline proposal for 25GbE RS/PCS/FEC (baden_3by_01b_0115) is the culmination of several earlier contributions which were presented and discussed at several 25GbE architecture ad-hoc calls and study group meetings.

- This presentation provides a reference to those contributions, and a brief overview of some of the key reasoning which led to the baseline proposal in baden_3by_01b_0115.

References

- Current baseline proposal:

  - http://www.ieee802.org/3/by/public/Jan15/baden_3by_01b_0115.pdf

- Previous related presentations:

  - http://www.ieee802.org/3/25GSG/public/adhoc/architecture/kim_100114_25GE_adhoc.pdf

  - http://www.ieee802.org/3/25GSG/public/adhoc/architecture/kim_100814_25GE_adhoc.pdf

  - http://www.ieee802.org/3/25GSG/public/adhoc/architecture/booth_102914_25GE_adhoc.pdf

  - http://www.ieee802.org/3/25GSG/public/Nov14/baden_25GE_01a_1114.pdf

  - http://www.ieee802.org/3/by/public/adhoc/architecture/baden_121714_25GE_adhoc.pdf
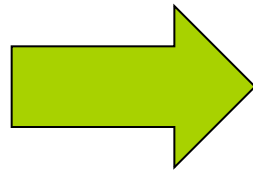
Starting point

CL46 — xMII (4b) — Clause 49 PCS — Clause 74 FEC — PMA — PMD

CL81 — xMII (8b) — Clause 82 PCS — Clause 74 FEC — PMA — PMD

CL81 — xMII (8b) — Clause 82 PCS — Clause 91 FEC — PMA — PMD

Clause 73 AN

10GBASE-xR | 40GBASE-xR4 | 100GBASE-xR4

Single-Lane Interfaces | Multi-Lane Interfaces

Starting point

Source: booth_102914_25GE_adhoc

xMII (4b)	xMII (?b)	xMII (8b)	xMII (8b)
Clause 49 PCS	PCS?	Clause 82 PCS	Clause 82 PCS
Clause 74 FEC	FEC?	Clause 74 FEC	Clause 91 FEC
PMA	PMA	PMA	PMA
PMD	PMD	PMD	PMD
Clause 73 AN			
10GBASE-xR	25GBASE-xR	40GBASE-xR4	100GBASE-xR4
Single Lane Interfaces		Four Lane Interfaces	
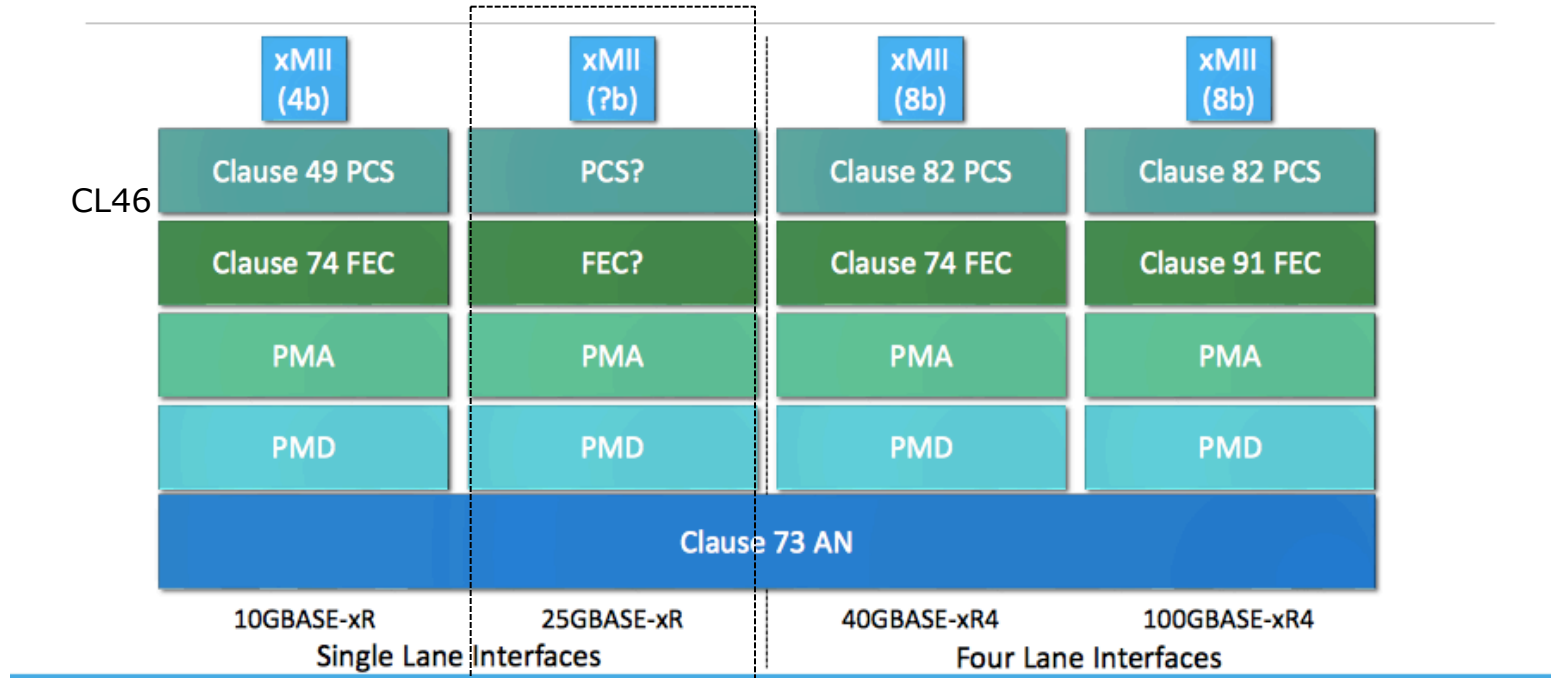
CL46

RS ?
PCS ?
FEC ?
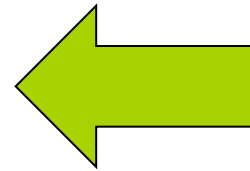
Move up from 10G ?
(CL46, CL49, CL74)

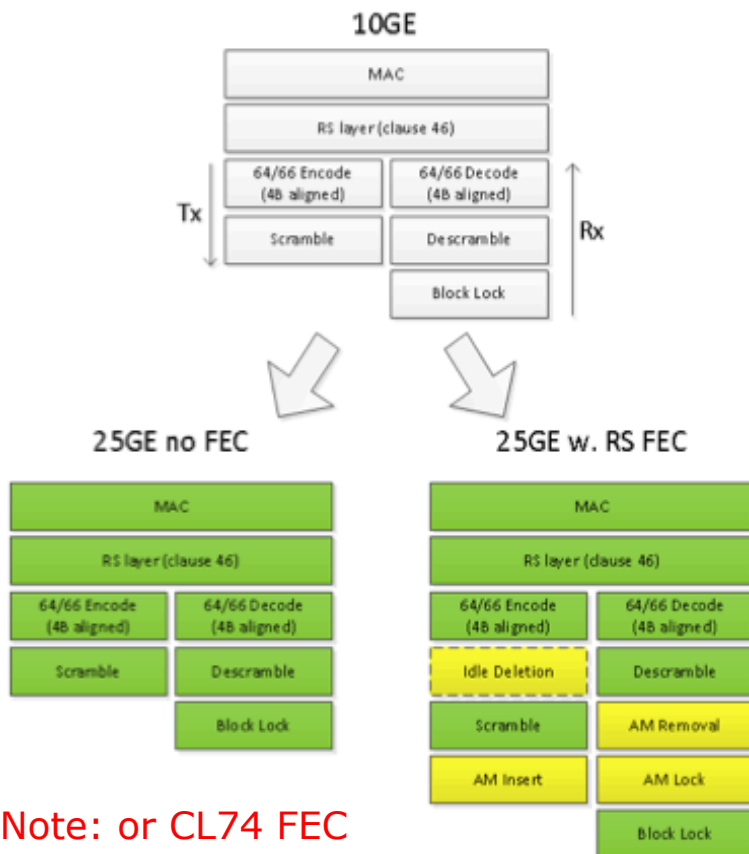Move down from 40G/100G ?
(CL82, CL74, CL91)

Application Review

- ## 10/25 NIC (lion's share of market ?)
  - primarily copper (and < 3m ?)
  - only ever needs to support single lane solution
  - likely SFP28 based
- ## 10/25/40 NIC
  - primarily copper
  - also needs to support multi-lane PCS (due to 40G)
  - likely QSFP based (but not a good fit for 10G/25G ?)
  - how likely is this due to connector mismatch (SFP v QSFP)
- ## 10/25/40/100 Switch
  - switch ports more likely to be multi-rate
  - therefore will have all PCS (single and multi-lane) and FEC versions available
  - likely QSFP28 based (/w breakout for 10/25 support)

PCS - Scaling up from 10G ?

25GE PCS using 10GE (CL49) building blocks

10GE

MAC
RS layer (clause 46)

64/66 Encode (4B aligned)	64/66 Decode (4B aligned)
Scramble	Descramble
	Block Lock

Tx → ← Rx

25GE no FEC

MAC
RS layer (clause 46)

64/66 Encode (4B aligned)	64/66 Decode (4B aligned)
Scramble	Descramble
	Block Lock

Note: or CL74 FEC

25GE w. RS FEC

MAC
RS layer (clause 46)

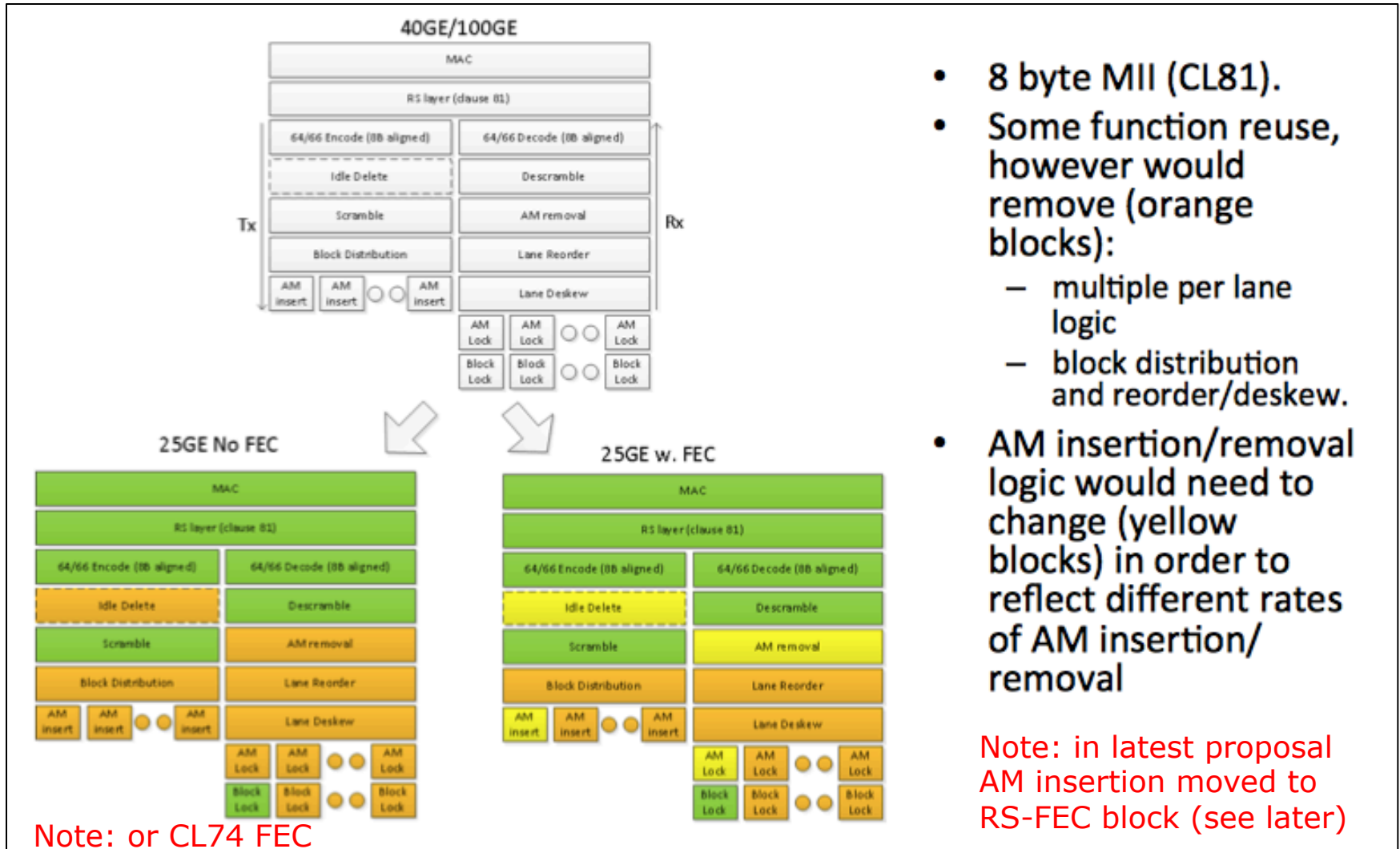| 64/66 Encode (4B aligned) | 64/66 Decode (4B aligned) |
Idle Deletion	Descramble
Scramble	AM Removal
AM Insert	AM Lock
	Block Lock

- 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)

Note: in latest proposal AM insertion moved to RS-FEC block (see later)

Simplest implementation for 10/25G NIC

PCS - Scaling down from 40/100G ?

- 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

Note: in latest proposal AM insertion moved to RS-FEC block (see later)

Note: or CL74 FEC

Need to change multi-lane arch to single lane arch

8

RS-FEC (CL91) Changes
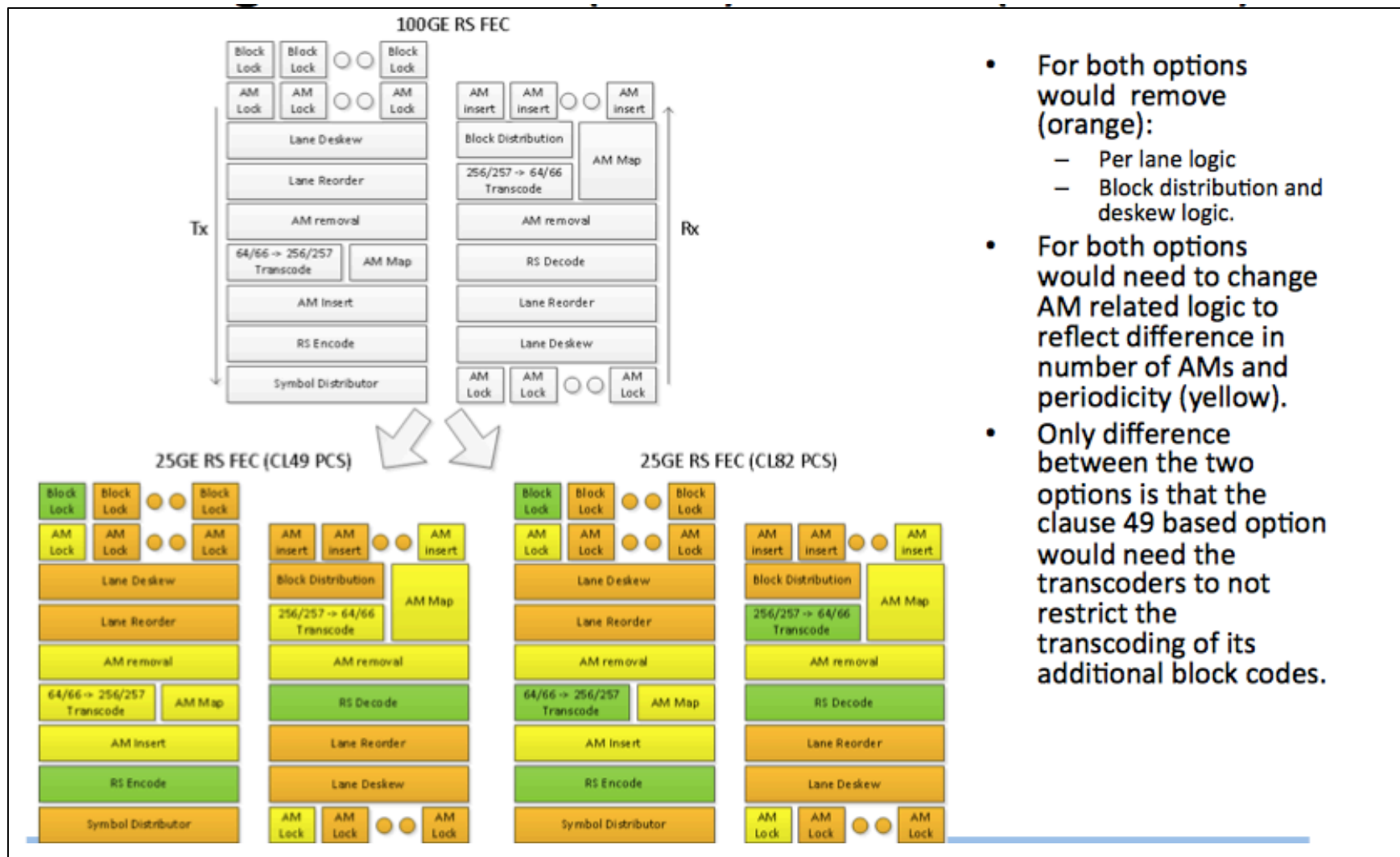
- For both options would remove (orange):
  - Per lane logic
  - Block distribution and deskew logic.
- For both options would need to change AM related logic to reflect difference in number of AMs and periodicity (yellow).
- Only difference between the two options is that the clause 49 based option would need the transcoders to not restrict the transcoding of its additional block codes.

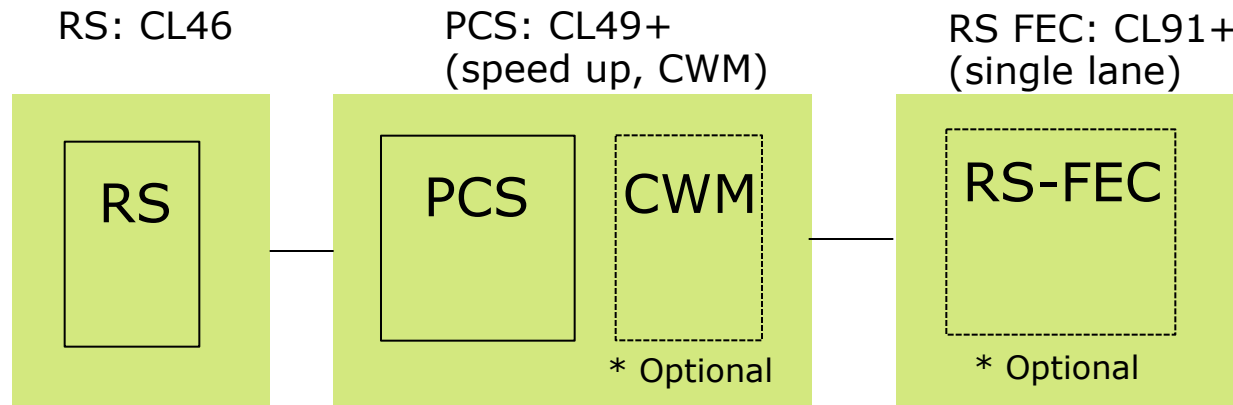Similar changes required independent of PCS starting point

Conclusions as of Dec 17 Ad Hoc

- Clause 49 is the better starting point for a 25GbE PCS.

  - even in the case where an AM (CWM) is required to be inserted to support an optional RS-FEC

  - clearly the simplest implementation for 10/25G NIC

  - one could argue that is also the simplest implementation for a multi-rate 10/25/40/100G switch port

- Changes are required to support a single lane Clause 91 based RS-FEC, irrespective of whether or not the 25GbE PCS is based on CL49 or CL82.

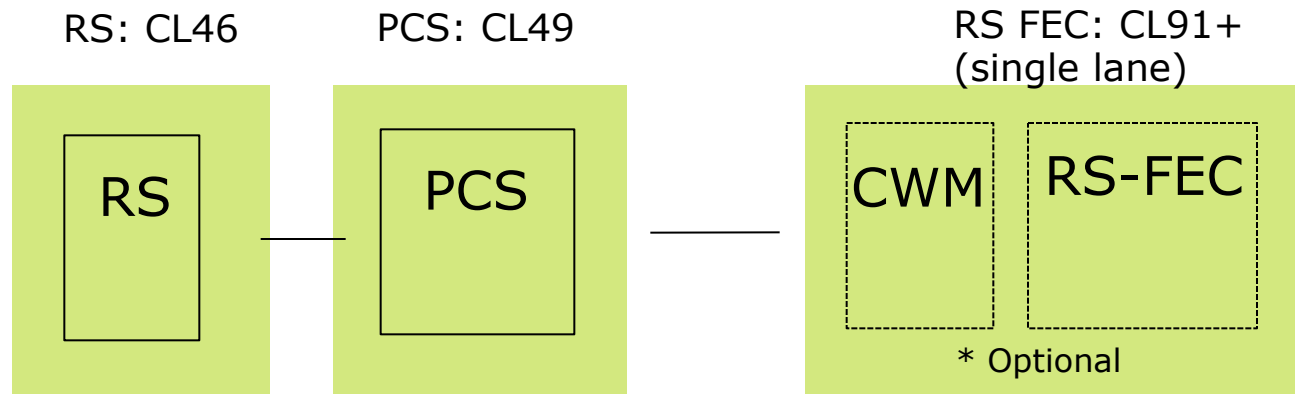 - magnitude of changes are equivalent in both cases

Change since Dec 17 Ad Hoc Call ?

- Moved AM (CWM) function that is needed to support RS-FEC, from being an optional mode in the PCS clause to being part of the RS-FEC clause.

Dec/2014

RS: CL46

PCS: CL49+
(speed up, CWM)

RS FEC: CL91+
(single lane)

RS — PCS | CWM
* Optional
— RS-FEC
* Optional

Jan/2015

RS: CL46

PCS: CL49

RS FEC: CL91+
(single lane)

RS — PCS — CWM | RS-FEC
* Optional

Note: Optional CL74 FEC not shown in interest of clarity.

Summary

- We believe the the latest baseline proposal (including the change in moving the CWM to the RS-FEC clause) is a good compromise to address all the different application needs for 25GbE.

- Best from an Architecture perspective
  - CWMs really belong with RS-FEC
  - Allows optional RS-FEC to be cleanly decoupled from PCS
  - Most flexibility on how CWMs are implemented (especially for chip designs which implement both PCS and FEC)

- Best from an Editorial perspective
  - Pretty much only one Clause (RS-FEC) has significant changes

- Best from a 'needs' perspective
  - A vast majority of companies will benefit from simply speeding up 10G KR related functions only