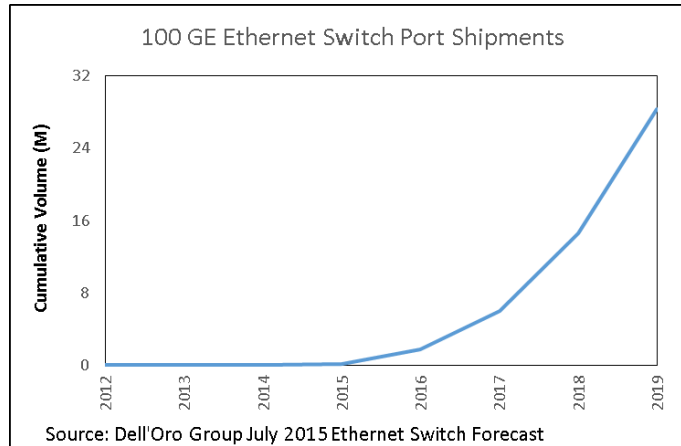# COMPATIBILITY CONSIDERATIONS FOR 50 AND 100G

Rob Stone

IEEE Study Group - 50G / NGOATH, Atlanta January 2016

- **Vasu Parthasarthy, Broadcom**

- **Tom Palkert, Molex**

- **Scott Somers, Molex**

- **Eric Baden, Broadcom**

- **Greg McSorley, Amphenol**

- **Brad Booth, Microsoft**

- **Both 100GE and 50GEc* are existing MAC rates**

- **There will be a large and growing installed base of 100GE (~ 15M cumulative in 2018)**
  - Data applies to 4 x 25 G interfaces used in switch applications
  - [These ports can typically be configured as 4 x 25GE, or 2 x 50GEc also]
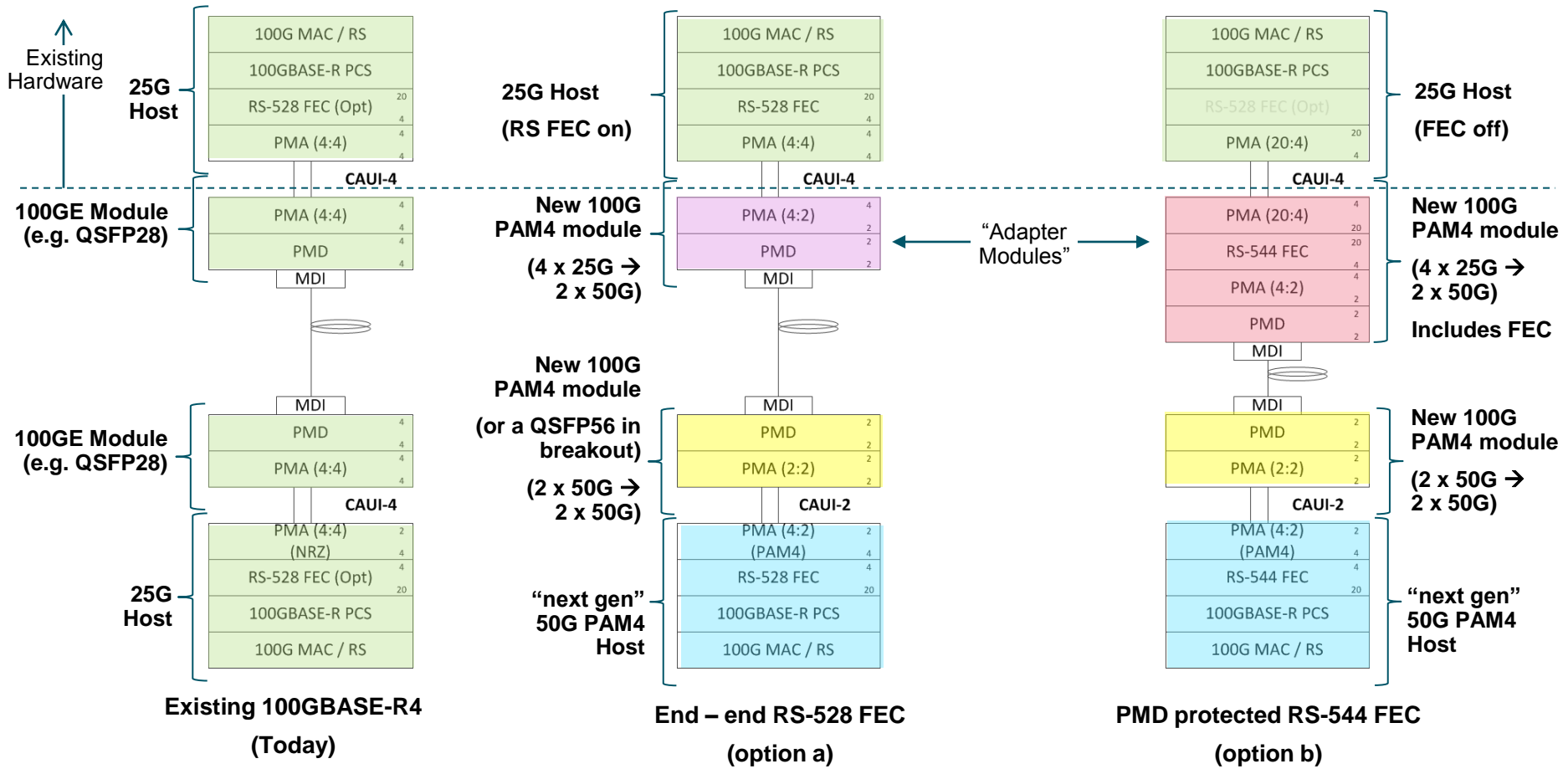
100 GE Ethernet Switch Port Shipments

Cumulative Volume (M)

32 · 24 · 16 · 8 · 0

2012 2013 2014 2015 2016 2017 2018 2019

Source: Dell'Oro Group July 2015 Ethernet Switch Forecast

Used with permission

- **Do we need to make consideration in 50 / NGOATH for connection of new 50G based PMDs to these "legacy" 100 & 50G ports based on 25Gb/s serdes?**

  * GEc used to denote 25 / 50G Consortium Specification

POSSIBLE GOALS FOR NEXT SET OF 50G PCS & PMDS

- **An architecture which permits straightforward connection of 50G / lane capable hosts to existing devices / technologies –**
  - 100GBASE-R4 PHYs
  - 50GEc (2 x 25G) PHYs – (non-IEEE variant - 25 / 50G Consortium Specification)

- **Offer low power and low latency where possible**

- **Plug and play for existing hardware**
  - Ideally an "adapter module" which enables use of both 50G / lane PMD and offers connection to next gen 50G capable host

- **Cost optimized**
  - Permits legacy designs to take advantage of new 50G based PMDs
  - Use appropriate FEC where possible to ease implementation challenges (applies to both optics and electrical PMDs)

- **Consider re-use where practical of existing work / developed IP**
  - FEC
  - MACs
  - PMAs
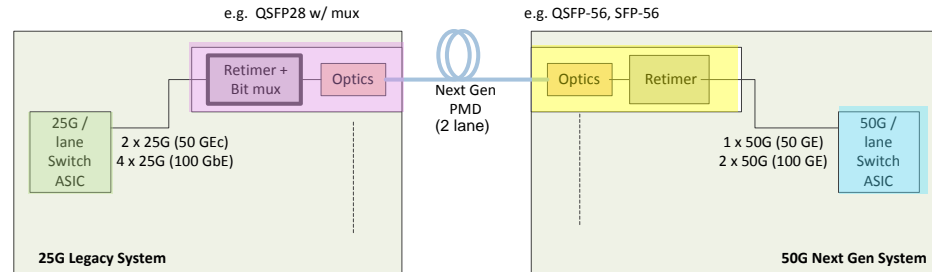 - Modules / management interfaces

Existing 100GBASE-R4
(Today)

End – end RS-528 FEC
(option a)

PMD protected RS-544 FEC
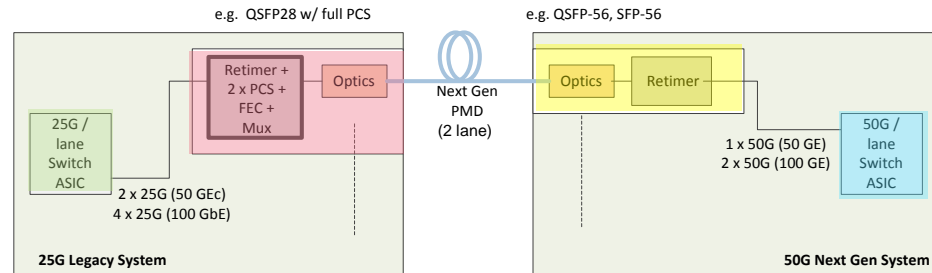**(option b)**

OBSERVATIONS (100G)

- **The simplest architecture relies on end-end FEC borrowed from 100GBASE-R4 (option a)**
  - This enables an QSFP28 with a simple PMA mux (4 x 25G → 2 x 50G lanes) – no change to CAUI-4 spec, or module mangement
  - Would enable plug and play upgrade to 100G over 2 lane PMD with *no hardware change on legacy designs*
  - Open question – does RS-528 provide adequate end – end gain for the 50G based PMDs and AUI?
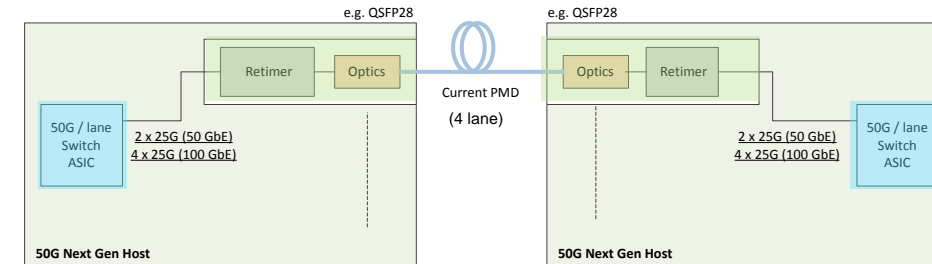 - Cons: no passive copper support, as requires the mux function in the module

e.g. QSFP28 w/ mux   e.g. QSFP-56, SFP-56

Retimer + Bit mux — Optics — Next Gen PMD (2 lane) — Optics — Retimer

25G / lane Switch ASIC — 2 x 25G (50 GEc) / 4 x 25G (100 GbE)

1 x 50G (50 GE) / 2 x 50G (100 GE) — 50G / lane Switch ASIC

25G Legacy System   50G Next Gen System

- **Alternative: add RS-544 FEC with appropriate PMAs to Module PHY or Gearbox on PCB (option b)**
  - **Drawback for Module:** Power envelope, BOM cost adder
  - **Drawback for PCB:** No longer a legacy hardware design
  - **Drawback for management** – SFF 8636 update / management for full PCS / FEC (would be required to be managed over I2C, or transition to MDIO and use CL45)

e.g. QSFP28 w/ full PCS   e.g. QSFP-56, SFP-56

Retimer + 2 x PCS + FEC + Mux — Optics — Next Gen PMD (2 lane) — Optics — Retimer

25G / lane Switch ASIC — 2 x 25G (50 GEc) / 4 x 25G (100 GbE)

1 x 50G (50 GE) / 2 x 50G (100 GE) — 50G / lane Switch ASIC

25G Legacy System   50G Next Gen System

- **"Do nothing" alternative is to not make allowance in 50G based projects for backwards compatibility**
  - Force 50G next gen designs to run in 25G "down speed" mode to connect to legacy
  - **Drawback:** Increases lane use by x 2 on next gen silicon, reduces switch radix, increases number of required switch stages to span a given network size

e.g. QSFP28   e.g. QSFP28

Retimer — Optics — Current PMD (4 lane) — Optics — Retimer

50G / lane Switch ASIC — 2 x 25G (50 GbE) / 4 x 25G (100 GbE)

2 x 25G (50 GbE) / 4 x 25G (100 GbE) — 50G / lane Switch ASIC

50G Next Gen Host   50G Next Gen Host

- **For 50 GE and 100 GE hosts using 50G / lane, should we consider an architecture which enables a streamlined connection to existing hosts based on 25G lane rate technology?**
  - Would enable use of 50G / lane PMDs for cost savings (2 x narrower # lanes)
  - Would enable next gen 50G host to operate at maximum bandwidth
  - Enables 25G and 50G based hosts to be used in the same environment without a significant penalization (hard to coordinate migration of all technologies to next lane speed FPGA, switch, NIC, …)

- **Simplest architecture (option a) would put a bit-mux in the module CDR, and requires that the AUI / PMD is protected with the existing end – end RS-528 FEC**
  - Is this adequate FEC gain *for a subset* of candidate PMDs?

- **An alternative view: Make no consideration for backwards compatibility with existing 25G based hosts in 50G / NGOATH projects**
  - i.e. require a 200GBASE-R4 port be operated down-speed at 100GBASE-R4 to connect to legacy 25G silicon
  - Drawback: this mode handicaps the next gen 50G host by requiring a down-speed of the AUI to 25G, with associated loss in radix / IO capacity and increase in network size

Thanks!