# Architectural Considerations for 50 GbE and NG 100 GbE
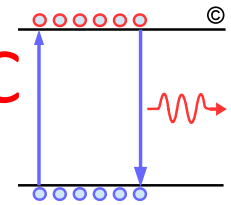
**Ali Ghiasi**

**Ghiasi Quantum LLC**

**IEEE 802.3cd Task Force  Meeting**

**Whistler**

**May 23,  2016**

# Compatibility, Synergy, and Interaction with FEC ©

- **In the task force need to investigate use of common FEC addressing all 50 GbE PMDs with likely choices being:**
  - RS-FEC (528,514) could satisfy 3 m Cu DAC but 30 dB backplane require RS-FEC (544,514)
    - See http://www.ieee802.org/3/cd/public/May16/ghiasi_3cd_02_0516.pdf
    - Market is served best by using single FEC RS(544,514) FEC which can meet Cu PMDs and optical PMDs
  - Compatibility and use case including operating with 25G MSA needs to considered
  - Need to support either bit mux if burst error is manageable otherwise symbol mux in support of 50GAUI-2/1
- **In the task force need to investigate possibly both RS-FEC (544,514) and (528,514) FEC:**
  - RS-FEC (528,514) could satisfy 3 m Cu DAC but 30 dB backplane require RS-FEC (544,514)
    - See http://www.ieee802.org/3/cd/public/May16/ghiasi_3cd_02_0516.pdf
  - Cu PMDs benefit from RS(544,514) but higher gain FEC could enable future MSA optics
  - In support of legacy 100 GbE PMDs also need to define CAUI-2 with RS(528,514)
  - Need to support either bit mux if burst error is manageable otherwise symbol mux in support of CAUI-4/2
- **Transition to 50G/lane optics may happen faster than migration to ASICs with 50G IO**
  - 50 GbE or NG 100 GbE implementation may take advantage of 400 GbE hardware which supports 16x25G electrical but 50G/lane or 100G/lane optics
  - To support flexible migration the 50 GbE PCS and NG 100 GbE PCSs should support respectively 50AUI-2/1 and CAUI-4/2 PMA Mux
  - If FEC performance is unacceptable with bit distributed PMA MUX, symbol distributed PMA is an acceptable alternative and preferable over implementing the full FEC in the PMA/PHY device
- **Application overlay should be key consideration to allow building common ports supporting overlay and breakout ports**
  - 1x400GbE, 2x200GbE, 4x100GbE, 8x50GbE.

# Possible 50 GbE Implementations
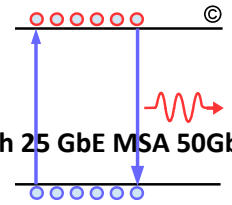
☐ **The key to supporting any existing 50 GbE or early implementation of 50 GbE is to support 50AUI-2 as well as support with 25 GbE MSA 50GbE over two lanes, what needs to be defined in 802.3cd**

    – Define 50GAUI-2 based on ½ of CAUI-4 under investigation

    – Define 50GAUI-2 based on 1/8 of CDAUI-16 see baseline in Li_3cd_01_0516.pdf

    – Define 50GAUI-1 based on 1/8 of CDAUI-8 see baseline in Li_3cd_01_0516.pdf.



A. Ghiasi

IEEE 802.3cd Task Force

3
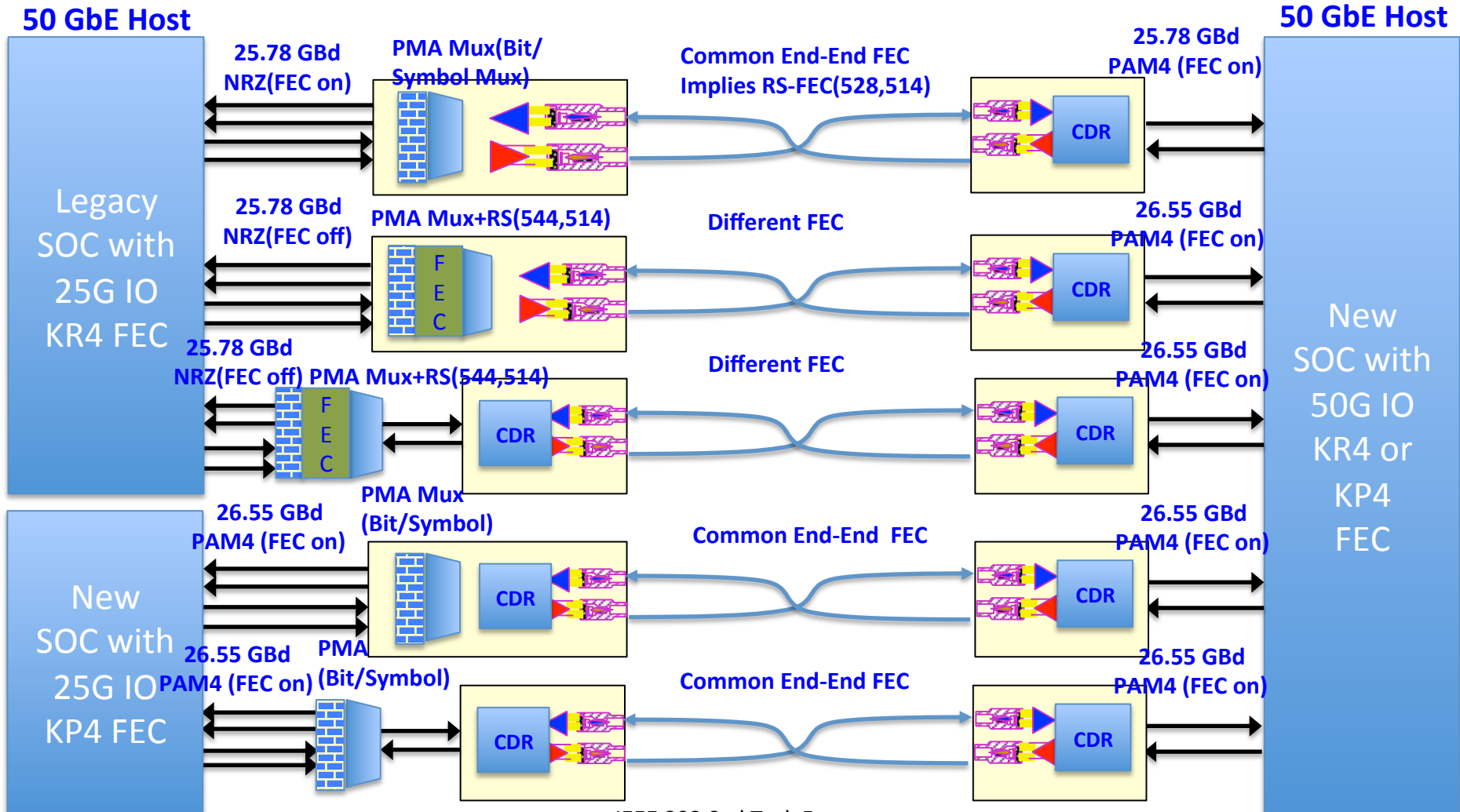
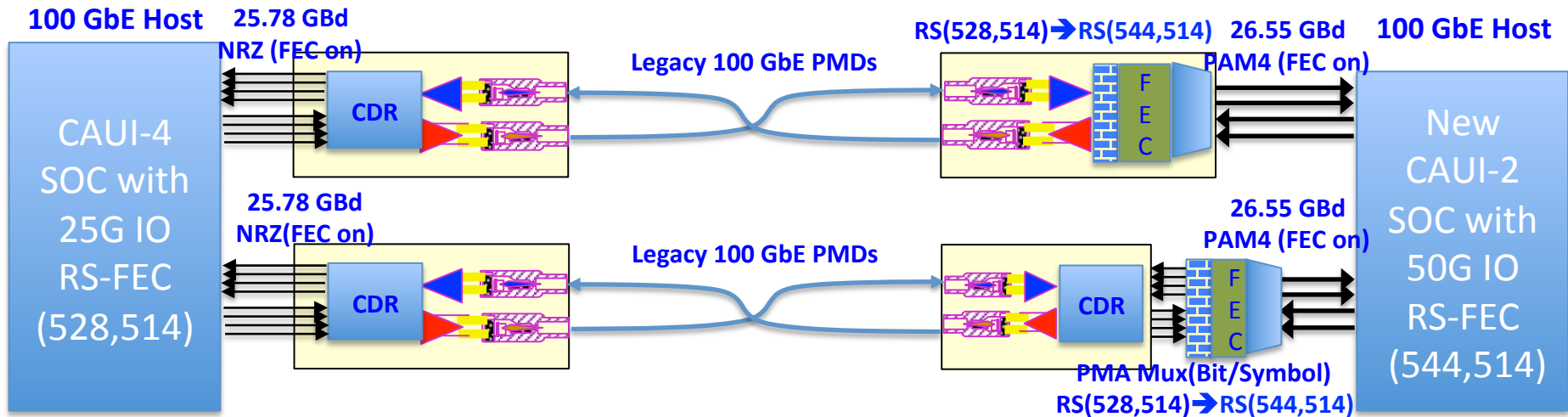# Legacy 100GbE PMDs Assuming Single KR4 FEC is Shared with CAUI-2
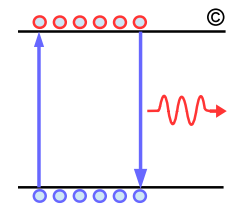
- ❑ **KR4 FEC with 20 PCS lanes offers highest level of backward compatibility to CL82 PCS and CL91 KR4-FEC**
- ❑ **Architecture shown below likely not be viable if RS(528,514) FEC gain is divided between CAUI-2 and legacy 100 GbE PMD that use the full RS(528,514) FEC gain**
  - To over come this limitation CAUI-2 would have to operate error free similar to CAUI-4 (1E-15) which maybe onerous
  - In engineered link application where optical link operate better than 5E-5 some of the FEC gain could be shared with CAUI-2
- ❑ **To support these use case need to define CAUI-2 with KR4 FEC**
  - Under investigation.
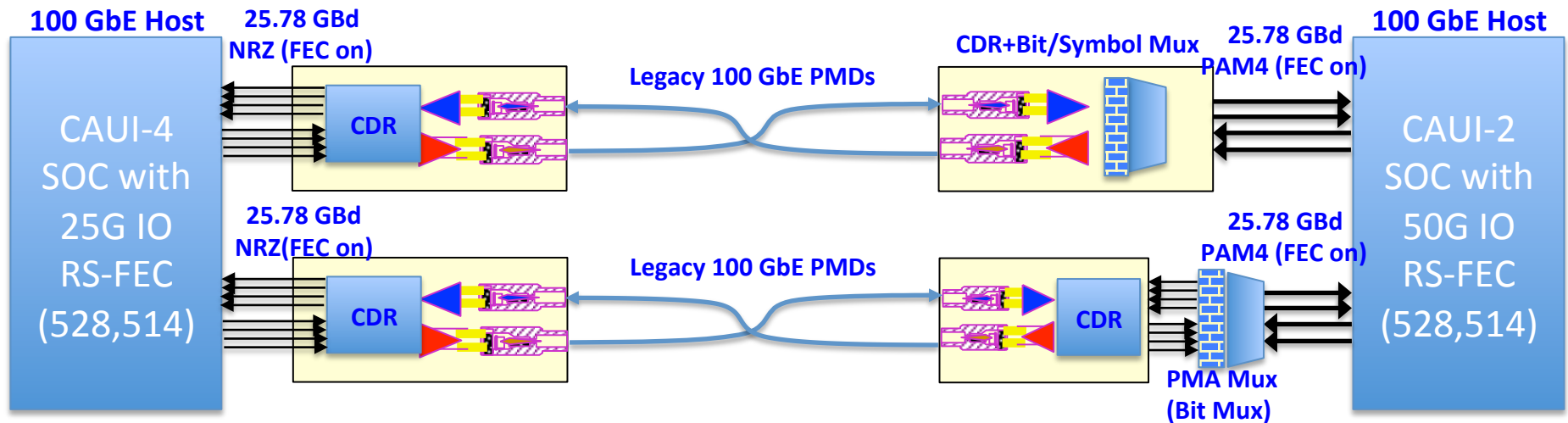
# 100 GbE Implementations of Legacy PMDs if CAUI-2 uses KP4 FEC
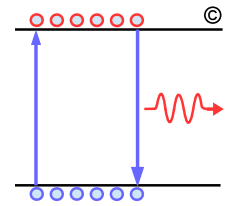
- ❑ **Need to balance the level of backward compatibility with overall synergy and burden**
  - – PMA-PMA+FEC device placed in module or on the line card can provide backward compatibility
- ❑ **To support this use case need to define CAUI-2 with RS(544,514) FEC**
  - – Currently there is no 100GbE optical PMD on the table so we don't have to define KP4 FEC for an application not under consideration in IEEE
  - – PMA-PMA extended sub layer in support of backplane and Cu DAC require defining CAUI-2 with RS(544,514) FEC based on ¼ of CDAUI-8
    - • See baseline in http://www.ieee802.org/3/cd/public/May16/li_3cd_01_0516.pdf



**100 GbE Host** — 25.78 GBd NRZ (FEC on) — Legacy 100 GbE PMDs — RS(528,514)➔RS(544,514) — 26.55 GBd PAM4 (FEC on) — **100 GbE Host**

CAUI-4 SOC with 25G IO RS-FEC (528,514)

CDR

25.78 GBd NRZ(FEC on)

Legacy 100 GbE PMDs

26.55 GBd PAM4 (FEC on)

CDR

New CAUI-2 SOC with 50G IO RS-FEC (544,514)

PMA Mux(Bit/Symbol) RS(528,514)➔RS(544,514)

\* 3 m Cu DAC could be satisfied with RS(528,514) but it is preferred to have common interface and FEC.
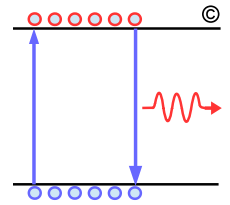
A. Ghiasi

IEEE 802.3cd Task Force

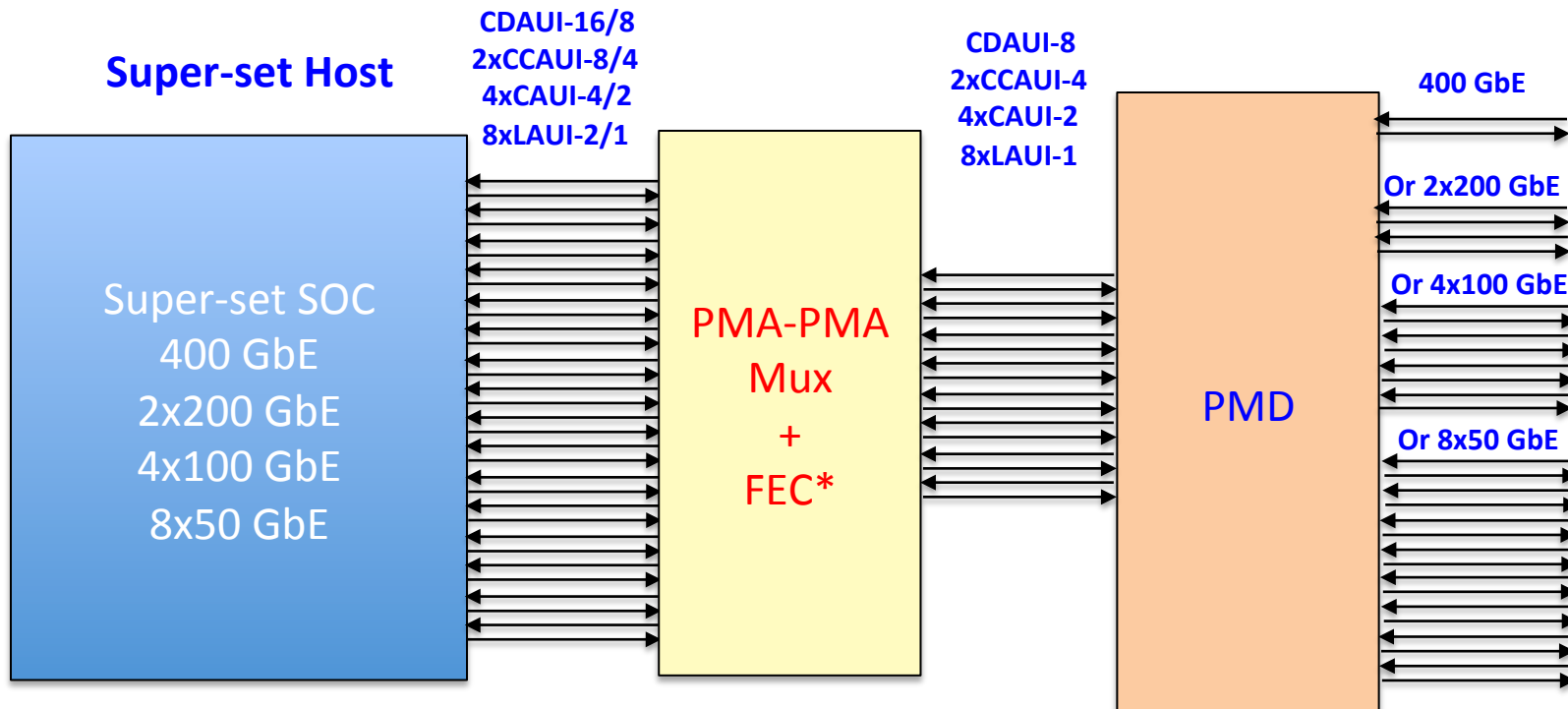# New 100 GbE PMDs Assuming KP4 FEC is Required
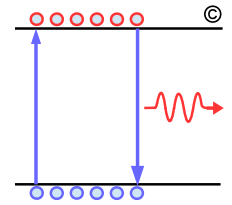
- ❑ **With no new 100 GbE PMD on the table IEEE 802.3cd does not have to consider these use cases**
  - – Defining CAUI-2 extended sub-layer in support of 100G-KR2/CR2 will also be a service to MSAs
    - • Define CAUI-4 based on 1/4 of CDAUI-16 see baseline in Li_3cd_01_0516.pdf
    - • Define CAUI-2 based on 1/4 of CDAUI-8 see baseline in Li_3cd_01_0516.pdf.

**100 GbE Host**

**CAUI-4 SOC with 25G IO RS-FEC (528,514)**

**New CAUI-4 SOC with 25G IO RS-FEC (544,514)**

**25.78 GBd NRZ(FEC off) FEC/PMA Bit/Symbol Mux**

**FEC**  **CDR**

**New 100 GbE PMDs**

**CDR**

**26.55 GBd PAM4 (FEC on)**

**100 GbE Host**

**25.78 GBd NRZ(FEC off) FEC/PMA Bit/Symbol Mux**

**FEC**

**New 100 GbE PMDs**

**CDR**

**26.55 GBd PAM4 (FEC on)**

**New CAUI-2 SOC with 50G IO RS-FEC (544,514)**

**26.55 GBd NRZ(FEC on) PMA Bit/Symbol Mux**

**CDR**

**New 100 GbE PMDs**

**CDR**

**26.55 GBd PAM4 (FEC on)**

**26.55 GBd NRZ(FEC on) PMA Bit/Symbol Mux**

**CDR**

**New 100 GbE PMDs**

**CDR**

**26.55 GBd PAM4 (FEC on)**

A. Ghiasi

IEEE 802.3cd Task Force
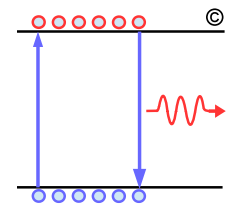
# Application Overlay

- **A key consideration of the 50 GbE and NG 100 GbE is application overlay of 200 G/ 400G ports in support of break out from a common port**
  - Supporting both 50GAUI-2/1 is required for 50 GbE application overlay shown below
    - CAUI-4 already exist
  - 25G/lane enables early implementations as well as transition from 25G MSA 50GbE mode.

**Super-set Host**

**CDAUI-16/8**
**2xCCAUI-8/4**
**4xCAUI-4/2**
**8xLAUI-2/1**

**CDAUI-8**
**2xCCAUI-4**
**4xCAUI-2**
**8xLAUI-1**

Super-set SOC
400 GbE
2x200 GbE
4x100 GbE
8x50 GbE

PMA-PMA
Mux
+
FEC*

PMD

**400 GbE**

**Or 2x200 GbE**

**Or 4x100 GbE**

**Or 8x50 GbE**

* In some implementation FEC may be part of the PMA-PMA otherwise in the super-set host.

# Example NIC/NPU with No Immediate Benefit from 50G I/O

❑ **50G I/O does not offer benefit in low port count applications**

  - Limited availability of 50G SerDes can be costly and/or not available

  - The ASIC may already have 50 GbE based on two lanes MSA

  - Not having the option of 50GAUI-2 may force a new ASIC development with costly 50G SerDes

  - The standard should offer flexibility and choice to either use 50GAUI-2 or 50GAUI-1.
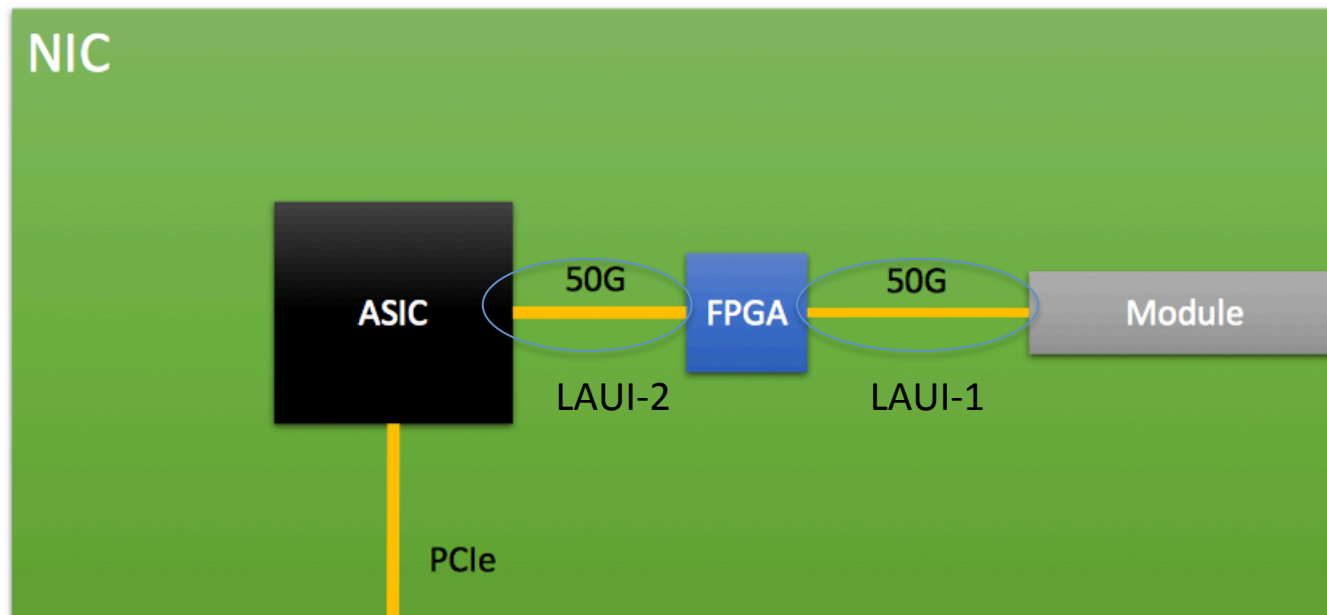


http://www.ieee802.org/3/50G/public/Mar16/booth_50GE_NGOATH_01a_0316.pdf

# Why 50GAUI-2 Is Needed
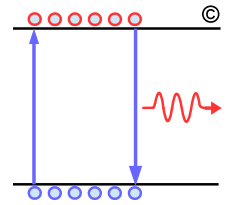
- **25 GbE 1$^{st}$ task force meeting was Jan 2015 where products (switches, NIC, Phys, CFP2, CFP4, and QSFP28) already available in the market place based on 25G/lane**
  - 25GbE was based on the definition of the 25G MSA which happened to leverage CL49 instead of the MLD PCS
- **Today the maturity of the 50G/lane IO is not at the same level as 25G IO was 2014/2015 time frame**
- **Following class of product do not benefit from migration to 50G/lane IO**
  - Core limited
  - NIC/NPU applications with just a few ports
- **Unless application can benefit from 50G IO (large switch) not having the option of the 50GAUI-2 forces the eco-system to higher risk and cost of 50G IO over commodity 25G IO!**

# Summary
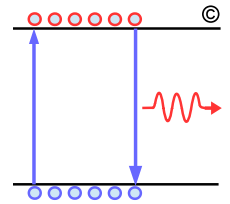
❑ **PCS/FEC options for 50 GbE**

– Speed up CL82 with 4 PCS lanes (12.5G) offers backward compatibility to the 25 GbE MSA (2x25G mode) and support 50GAUI-2/1

– Need to define 50GAUI-2/1 with RS(544,514) FEC

- Where 50GAUI-2 is 1/8 of the CDAUI-16
- Where 50GAUI-1 is 1/8 of the CDAUI-8
- Defining 50GAUI-2 being ½ CAUI-4 need further study

❑ **PCS/FEC options for NG 100 GbE**

– Use CL82 100GbE PCS base on 5G PCS lane offering backward compatibility and supports CAUI-4/CAUI-2

– Need to define CAUI-4/2 with RS(544,514) FEC

- Where CAUI-4 is 1/4 of the CDAUI-16
- Where CAUI-4 is 1/4 of the CDAUI-8
- Defining CAUI-4 with RS(528,514) needs further study

❑ **With 25G MSA 50 GbE mode and 100 GbE already in the market place IEEE needs to consider compatibility scenarios illustrated here as well as support for PMA bit/symbol mux to ease transition.**