Existing DSP Can Support 100 m Reach Objective

Ali Ghiasi – Ghiasi Quantum

IEEE 802.3dB Taskforce Meeting

Sept 17, 2020

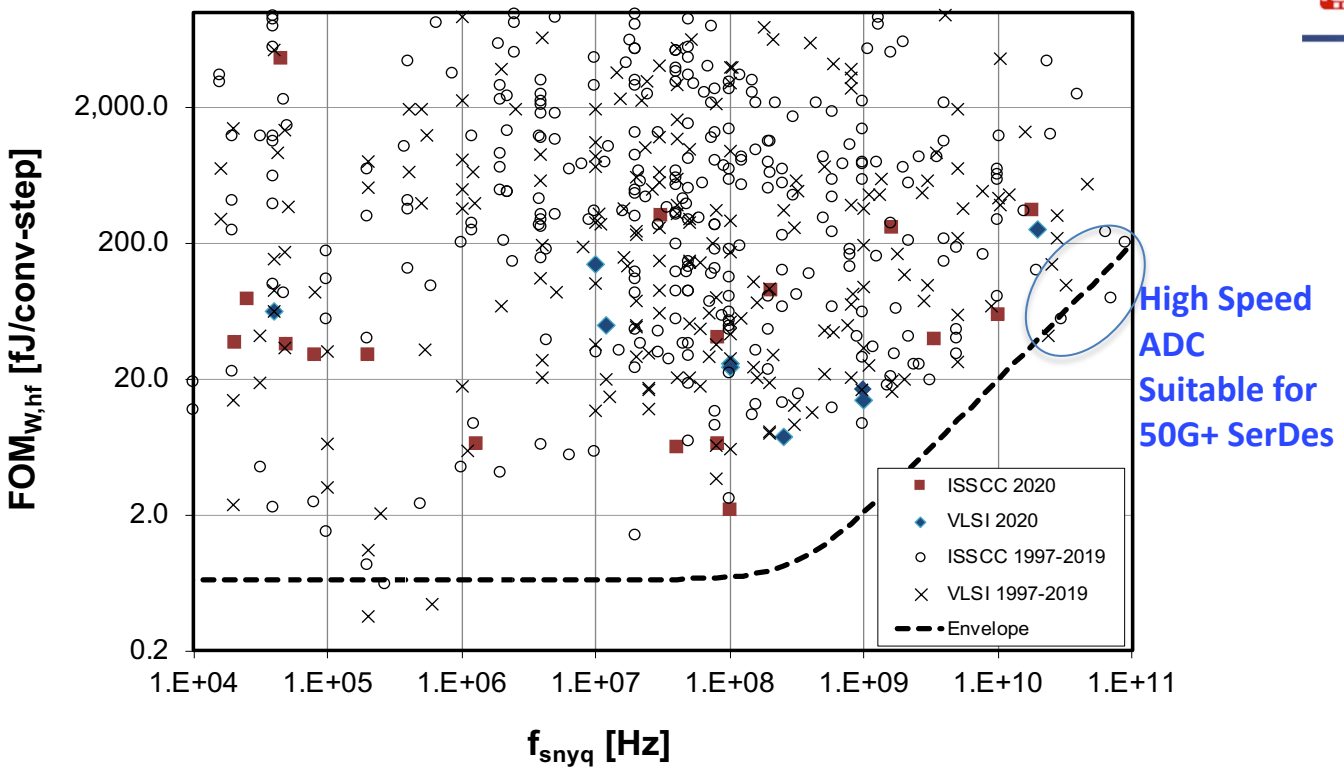
Overview

- ADC evolution
- Example of DSP SerDes/Phy
- Comparisons of DSP vs analog CDR
- CDR latency vs the network latency
- Reference equalizer to support 100 m PMD
- 802.3db doesn’t have the luxury of time
- Summary.

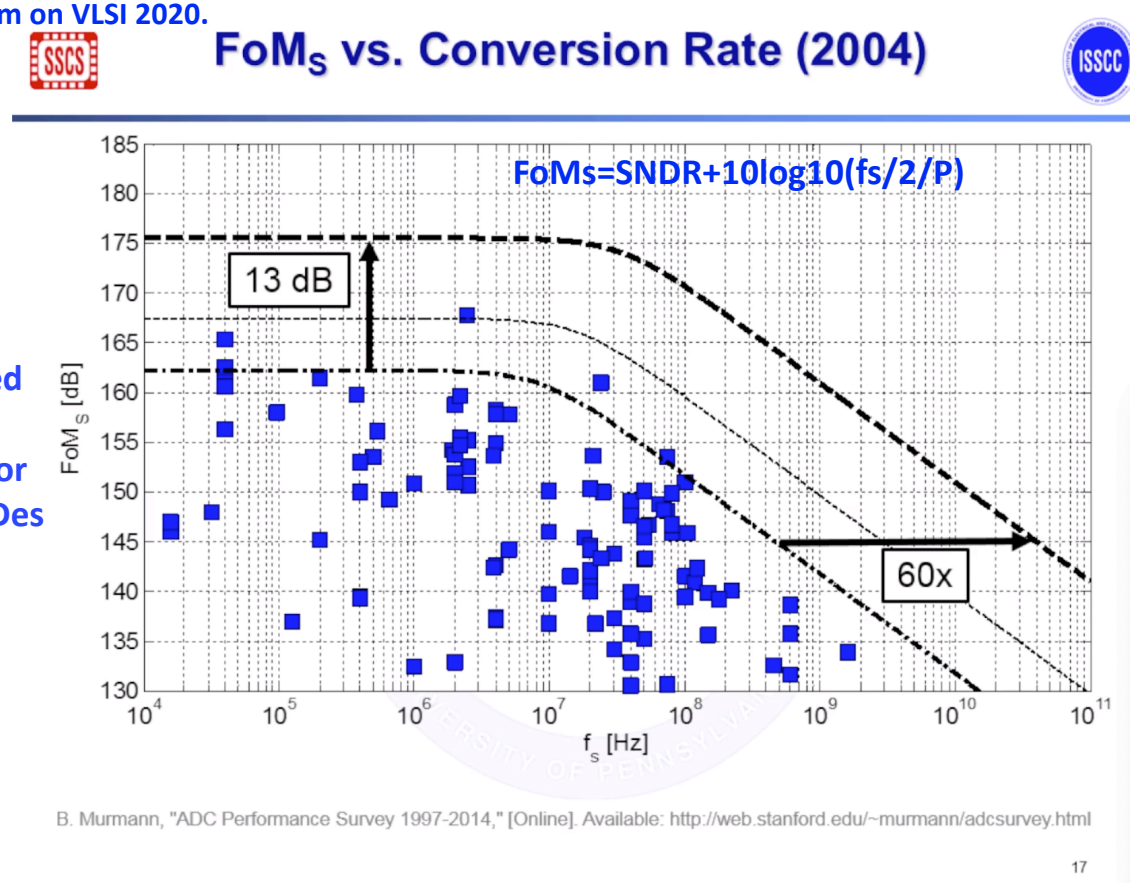
ADC Based SerDes Power-Performance Crossed Over with 16 nm CMOS

- **ADC have been improving at at rate of 2x/14 months which is faster than Moore's Law**
  - The drawback of DSP SerDes has been ADC power dissipation
  - 112 Gb/s 7 bits SAR ADC capable of 36 dB Cu reported PD is <200* mW 7 nm CMOS (include clocking power)
    - A 6 bits ADC reduces the ADC power by ~1/2 would be sufficient optical channels.

*** H. Lin, et. al., A 4×112 Gb/s ADC-DSP Based Multi-standard Receiver in 7nm FinFET, Symposium on VLSI 2020.**



https://web.stanford.edu/~murmann/adcsurvey.html



Overview of DSP Transceiver

- **Example of an older 50G-PAM4 transceiver in 28 nm**
  - Today's XCVR's in 7 nm and the digital blocks such as ADC and DSP area shrinks by ~4x and power reduced by ~3x
  - This design had 10 Taps FFE + 1 DFE which is a common DSP implementation

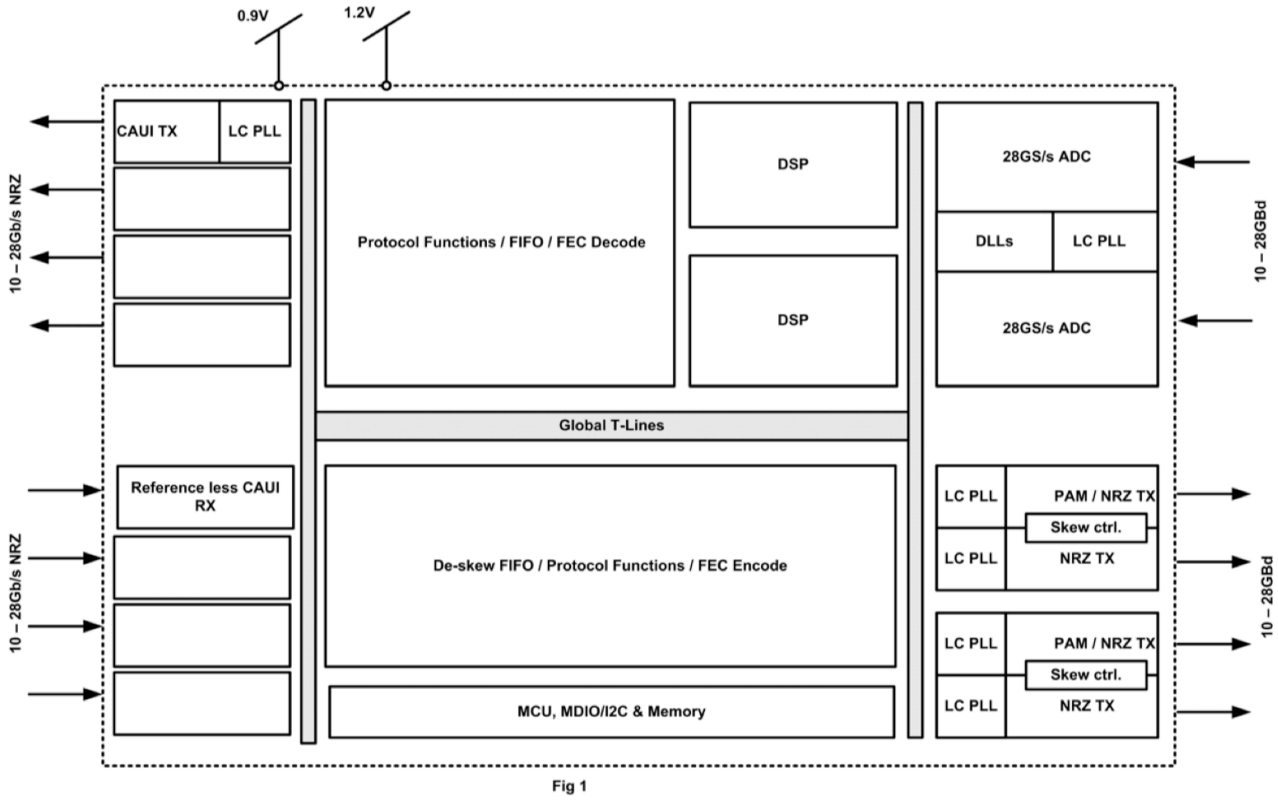


Fig 1

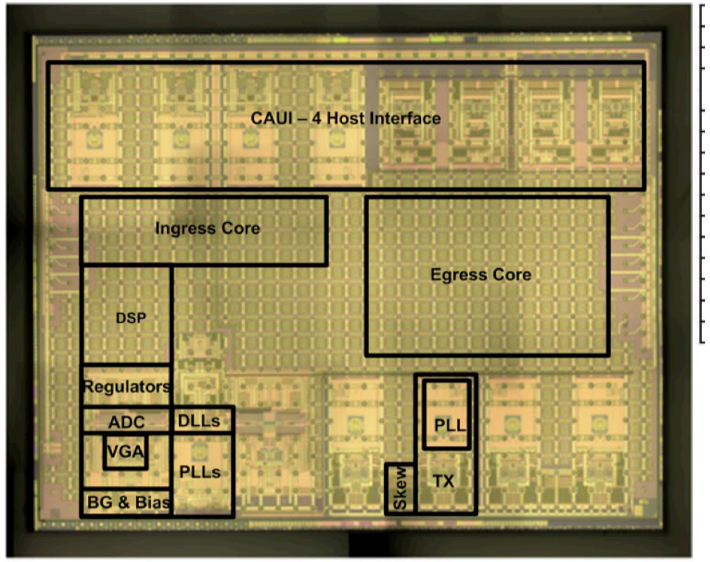


Fig 7

K. Gopalakrishnan, ISSCC 2016

DSP vs Analog CDR for Optics

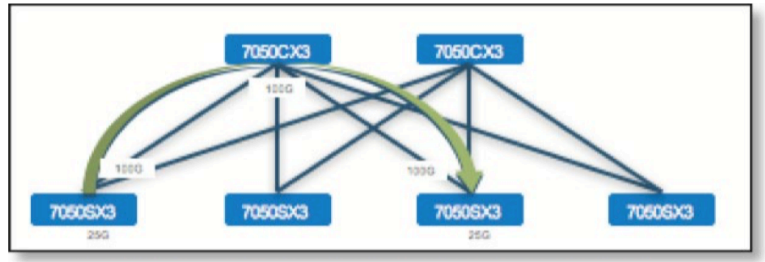
- **DSP CDR based on advance CMOS for PAM4 Signaling**
  - Advantages
    - More powerful equalization come at very little cost
    - Time to market
    - More robust blind adaptation and optimization to dynamic effect
    - Extensive diagnostics, self test/pattern Gen, and loop back
    - Lower power/tap
    - Smaller die size (lower die cost)
    - Yield
  - Disadvantages
    - Mask set more costly
- **Analog CDR based on advance BiCMOS for PAM4 Signaling**
  - Advantages
    - Lower mask cost
    - Lower latency (my estimate is 10 ns for analog vs 40 ns for DSP for 100G CDR, in the Open Eye MSA organized by OIDA it was reported 40 ns for analog CDR and 140 ns for DSP likely skewed by the measurement equipment and use of older DSP)
  - Disadvantages
    - More than 5-7 taps FFE are not technicality feasible
    - Very challenging to do Baudrate DFE
    - A CDR with 5-7 taps FFE will end up with higher power today than the DSP equivalent in 7 nm CMOS
    - Takes many design iterations for working product
    - Blind adaptation and dynamic optimization is problematic.

How Does the CDR Latency Compare to Network Latency?

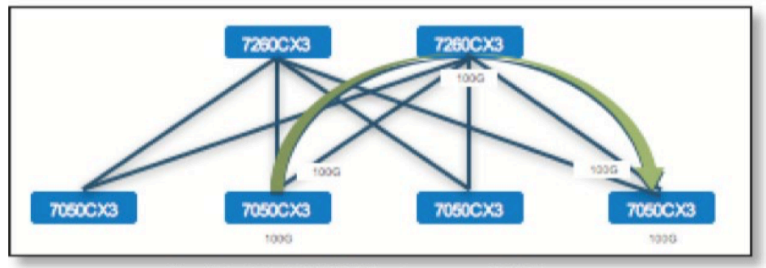
- **A very popular 3.2 Tb (32x100GbE) switch based on single ASIC from Arista has port-port latency of 800 ns**
  - But no network is built with just one switch
- **What matter is the latency for 3 and 4 tiers network than single ASIC latency**
  - Port-port latency is 3-4 μs if one uses 25GbE breakout for 2 tier network
  - If one implements end-end 100G cut-through then latency is 2-2.2 μs for 2 tier network
  - A more realistic network with 3 tiers estimated latency would be ~8-10 μs excluding fiber delay
    - A modest 500 m end-end fiber media will add 2.5 μs
- **A realistic 3 stage Ethernet network will have ~10 μs end-end latency**
  - A 3-stage network with 10 ns/CDR will add total of 40 ns or 0.4% to the latency
  - A 3-stage network with 40 ns/CDR will add total of 160 ns or 1.6% to the latency
- **CDR latency of 10's ns is in the noise in comparisons to typical Ethernet network end-end latency!**

Arista 3.2 1RU Switch

https://www.arista.com/assets/data/pdf/Datasheets/7050X3-Series-Quick-Look.pdf



25G/100G Leaf Spine Port to Port Latency ~3-4us



End to End 100G Latency ~2-2.2us

Figure 24: 7050X3 server to server latency

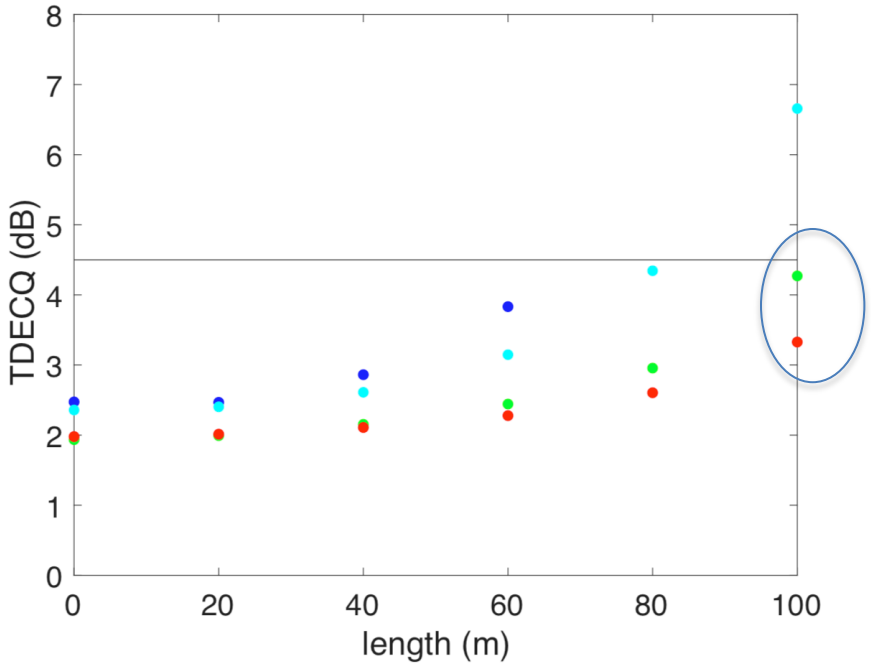
https://www.arista.com/assets/data/pdf/Whitepapers/7050X3_Architecture_WP.pdf

What Should the 802.3db Reference Equalizer be?

- **802.3bs initial choose reference equalizer with FFE**
  - In 2015 time frame the task force choose 5 taps T/2 equalizer
  - Given that 5 taps T/2 was insufficient in span to equalize the channels the task force in 2106 time frame changed the equalizer to 5 taps T spaced
- **The selection of taps FFE was driven by**
  - Define the shortest equalizer necessary that also allow lower power analog FFE implementation
  - Since 2015 CMOS node has evolved from 28 nm to 7 nm where DSP implementation offer comparable power for much greater capabilities
  - Analog implementation of 5T FFE has proved more challenging to implement as product compare to academic result reported
    - Due to these challenges OpenEye MSA was started to define receivers without FFE/DFE equalizations
- **Several leading CMOS SerDes suppliers offer DSP CDR for 100G PAM4 with significantly more capability than 5T FFE**
  - Every 100G-single λ optics in development or shipping uses one of these DSP chips or cores
  - These cores likely will have 15+ FFE taps and 1 DFE taps
  - These more chips are available from multiple suppliers and there is no extra cost or power to use their full capabilities beyond 5T FFE!

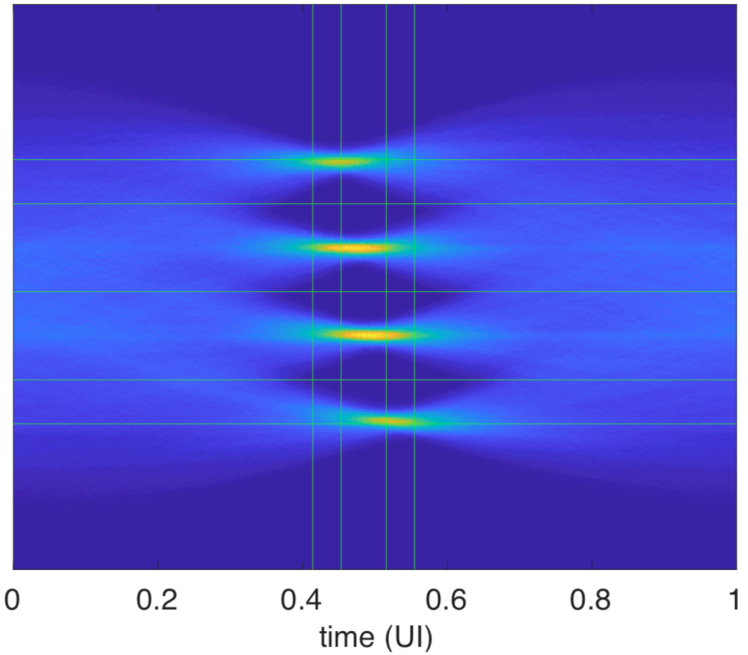
Minimum Equalizer to Support 100 m on OM4

- **Ingham results indicate that at least 9 tap FFE is required to support 100 m reach on OM4 fiber**
  - Given the steep increase in TDECQ from 80 to 100 m (~3.3 dB to ~4.3 dB) the real equalizer would need 12-15 taps FFE to have some margin!



TDECQ versus OM4 link length for 5-tap Rx FFE (blue), 7-tap Rx FFE (cyan), 9-tap Rx FFE (green) and 23-tap Rx FFE (red)

If TDECQ is not measurable at a particular length, then a data point is not shown



Eye diagram after 100 m OM4 with 9-tap Rx FFE

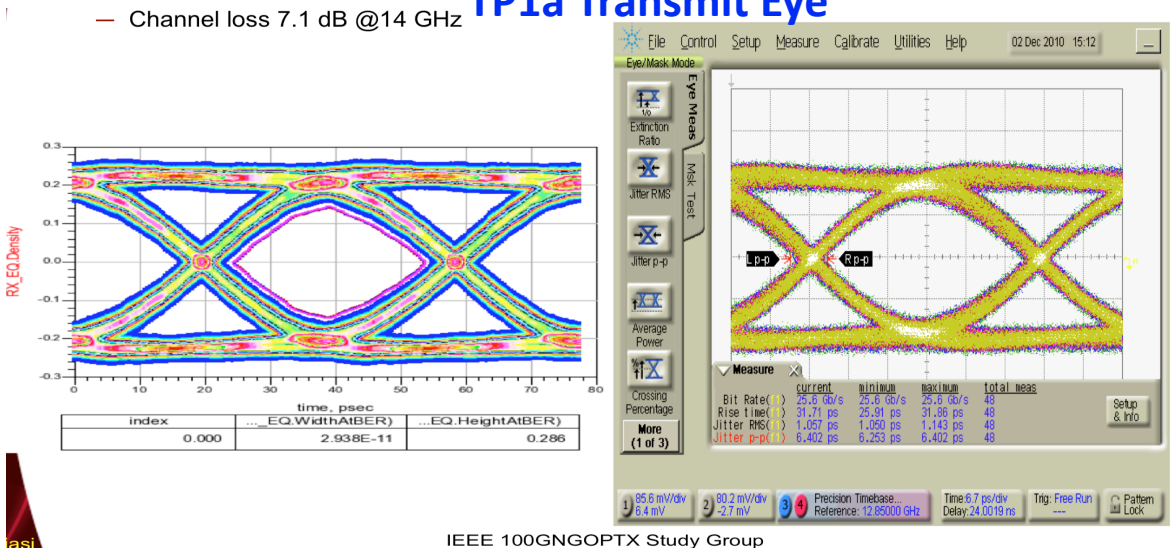
Rx noise not depicted; green lines show outer levels, thresholds and optimally-positioned time windows for TDECQ

Feasibility of Unretime/Direct Drive

- **Unretimed cPPI-4 with 7 dB loss proposed in 802.3cm in 2011 as alternate to retimed CAUI-4**
  - Unretime/direct drive are not new concepts
    - https://www.ieee802.org/3/100GNGOPTX/public/nov11/ghiasi_02_1111_NG100GOPTX.pdf
  - In 2012 I published Feasibility of unretime with 25.78 GBd 100GBASE-LR4
    - See Ghiasi and et. al., OFC 2012, OW1J.2
    - See https://www.osapublishing.org/abstract.cfm?uri=OFC-2012-OW1J.2
  - In 2013 I published Feasibility of unretime with 25.78 GBd VCSELs
    - See Ghiasi and et. al., OFC 2013, OM2H.3
  - Simulations and measurements included both electrical and optical channels.

TP1a Transmit Eye

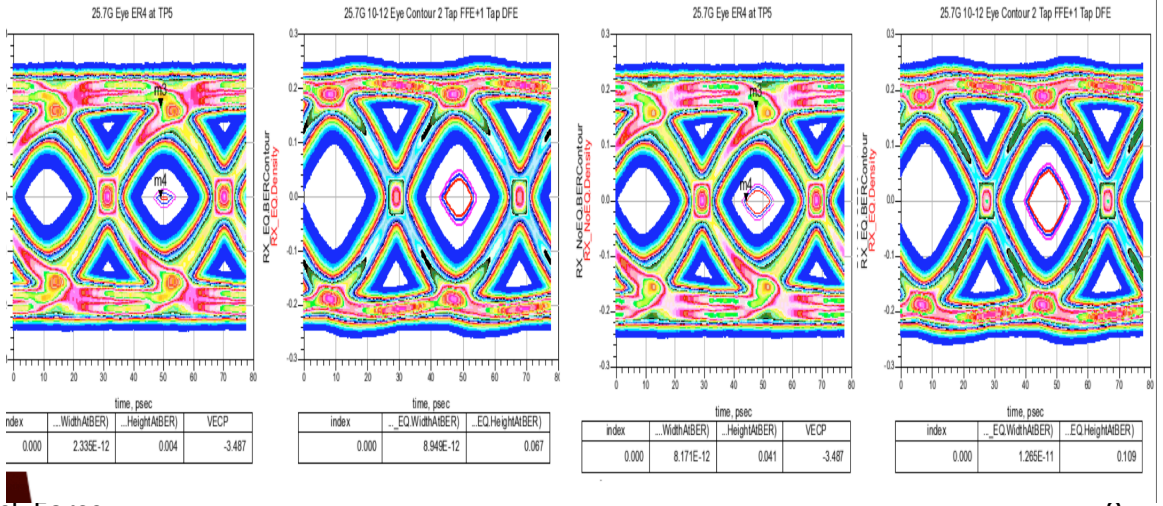
– Channel loss 7.1 dB @14 GHz



TP5 Receive Electrical-Optical-Electrical Eye

- **Results at TP5 with 0.3 and 0.4 UI at sampling point**

Increasing sampling eye opening by 0.1 UI require ~ 2 dB sensitivity improvement



Feasibility of Unretime/Direct Drive, cont.

- **Unretime/direct drive at 25.78 GBd is not a new concept**
  - IEEE 802.3bm end up going with retimed interface but unretime did have substantial support in the industry
  - Both MMF and SMF link were demonstrated
- **What are the key challenges in unretime/direct drive**
  - The link consist of TX electrical channel + Optical link + RX electrical link
  - VCSELs are nonlinear, require non-linear drivers and receive equalizer may need to be Voltera
  - The simulation or measurement must include electrical and optical interactions
  - Electrical TX jitters behave as if the optical link Baudrate is increased
  - Interop/testability with existing optical PMDs can be challenging
- **What can we do now**
  - With new improved CTLE @ 25.78 GBd NRZ likely PPI loss can be increased to 10 dB just like CAUI-4
  - PAM4 is extremely challenging for unretime/direct drive due to reflections, in the 802.3ck project the 4T DFE may not be sufficient to equalize a 2” channel
  - Using 100G connector likely we can develop 50G unretime/direct drive and if one restrict the loss between ~5 to 8 dB
  - 100G Unretime/direct drive is beyond technical feasibility when a simple point-point 100G-AUI1 link require more than 4T DFE!

The 802.3db Doesn't Have the Luxury of Time

- **SMF products based on 100G-DR/400G-DR4 are now shipping in volume**
  - Unless the db task force stays on schedule many of customers planning upgrade will choose SMF over MMF
  - 802.3ck CFI was on Nov. 2017 and after about 3 years we have just produced D1.3
  - Unretime/direct drive at 53 GBd PAM4 is significantly more challenging (if even possible) than 802.3ck
  - The task force need to focus on technical feasibility and march to D1.0 with timeline provided by Mr. Lingle preferably get to D1.0 no later than March 2021!
    - .3db PAR: May-2022 standard
    - .3db A: May-2022 standard. Oct-2020 preliminary baseline; Nov-2021 baseline, Jan-2021 D1.0
    - .3db B: Jul-2022 standard. Nov-2020 preliminary baseline; Jan-2021 baseline, Mar-2021 D1.0
    - .3db C: Jan-2023 standard. Nov-2020 preliminary baseline; Mar-2021 baseline, Jul-2021 D1.0
      - 2 cycles between prelim and adopted baseline; 2 cycles between baseline and D1.0; D2.2 added

	.3cm	.3db	.3db - PAR	.3db - A	.3db - B	.3db - C	.3ck	.3cu
1st TF meeting	May-18	Jun-20	Jun-20	Jun-20	Jun-20	Jun-20	May-18	May-19
Objectives adopted (WG)	May-18	Jun-20		Jun-20	Jun-20	Jun-20		
Objectives updated	N/A	?		?	?	?		
Baselines preliminary	May-18	?		Oct-20	Nov-20	Nov-20		
Baselines adopted	Jul-18	?		Nov-20	Jan-21	Mar-21	Nov-19	Sep-19
D1.0	Sep-18			Jan-21	Mar-21	Jul-21	Jan-20	Sep-19
Standard	Jan-20		May-22	May-22	Jul-22	Jan-23	Sep-21	Dec-20
1st SG to Standard	2y		2y, 4m	2y, 4m	2y, 6m	3y	3y, 8m *	1y, 11m
1st TF to Standard	1y, 8m		1y, 11m	1y, 11m	2y, 1m	2y, 7m	3y, 4m *	1y, 7m

Summary

- **Given the Ingham results the 100 m on OM4 can be supported with just 9 taps FFE**
  - The good news for the task force is that it is a solved problem as most DSP CDR have more than 15 taps
  - These CMOS CDR/PHY chips are shipping in volume enabling 100G/400G single λ optics
  - Our current objective of at least 50 m can either be changed to 100 m or could be met with 100 m reach PMD
- **The DSP equalizers can operate with nominal fix TX FIR and will outperform shorter span receive equalizer having an adaptive transmitter as Castro suggest**
  - VCSEL property, EMB effects, DMD effects, and channel lengths can vary and require initial adaptation and further adaptation during operation
  - Receive equalizer performs continuous adaptation and don't require any change to PCS
- **The good news for the task force is that adaptive equalizer with at least 15 tap FFE are shipping in volume**
  - By leveraging these more capable proven products 100G-SR optics can start sampling in 12 months
  - No need to re-invent the wheel
  - Touching Ethernet PCS is a non-starter
- **Given that timeline of the db task force to get to D1.0 no later than Jan-2021 the focus should be developing optical PMDs instead dabbling in the unretime/direct drive interfaces**
  - Unretime/direct drive at 25.78 GBd NRZ technical feasibility was demonstrated in 2011
  - Unretime/direct drive at 26.55 GBd PAM4 may have technical feasibility to merit further investigation
  - The 25G NRZ and 50G PAM4 unretime/direct drive with some technical feasibility are not relevant to db project
- **By the time standard is published CMOS CDR's will have 30-50% lower power with Moore's Law evolution!**