Ethernet in Design (100Gb/s Lane Rate)

Moderator:

Kae Dube, Ethernet Operations Manager, UNH-InterOperability Lab

Presenters: Nathan Tracy, Technologist, System Architecture Team, TE Connectivity Adithya Muralidharan, Staff Applications Engineer, Intel John Calvin, Senior Strategic Planner, Keysight Pavel Zivny, Domain Expert, Tektronix



www.ethernetalliance.org

Presentation Disclaimer

The views expressed in this panel presentation are those of the presenters and not of the Ethernet Alliance.



3-Part Series

on demand -> Webinar 1: Ethernet in the Field (50Gb/s Lane Rate)

Today Webinar 2: Ethernet in Design (100Gb/s Lane Rate) Webinar 3: Ethernet in the Future (200Gb/s Lane Rate)



www.ethernetalliance.org

About Ethernet Alliance

Kae Dube, Ethernet Operations Manager, UNH-InterOperability Lab



www.ethernetalliance.org

The Ethernet Alliance

Global Community of End Users, System Vendors, Component Suppliers & Academia

Our Mission

- To promote industry awareness, acceptance and advancement of technology and products based on, or dependent upon, both existing and emerging IEEE 802 Ethernet standards and their management.
- To accelerate industry adoption and remove barriers to market entry by providing a cohesive, market-responsive, industry voice.
- Provide resources to establish and **demonstrate multi-vendor** interoperability.





Ethernet Alliance Strategy

Expanding the Ethernet Ecosystem, Supporting Ethernet Development

• Facilitate interoperability testing & assurance

- Industry Plug fests supporting member and technology initiatives
- PoE Certification Program
- Global outreach and collaborative interaction with other industry organizations
- Worldwide Membership
- Multiple SIGs, applications and MSAs
- Industry consensus building
- Thought Leadership
- EA-hosted Technology Exploration Forums (TEFs)
- Technology and standards incubation
- Promotion of Ethernet
- Media and industry analysts outreach
- Education
- Marketing (trade shows & panel presentations, white papers, blogs & social media)

2023 Ethernet Roadmap

- Digital copies & graphics published via Alliance's website
- Print copies available at OFC 2023 and other events this year
- Included in the "Ethernet Alliance in the Box" for members
- Digital version regularly updated to capture latest advancements







How Ethernet is Enabling the Next Generation; 100Gbps Lane Rate

Nathan Tracy, Technologist, System Architecture Team, TE Connectivity



ethernet alliance

www.ethernetalliance.org



Ethernet 100Gbps Lane Rate Enablers

IEEE's 802.3 Ethernet Standard has developed a number of new Clauses to enable and leverage 100Gbps signaling per lane rate to meet next generation networking needs:

802.3ck: Establishes the electrical specifications and options (more details to follow)

802.3ct: Establishes the DWDM optical specifications for applications with reach of at least 80km

802.3cu: Establishes the single mode 100Gbps wavelength optical specifications and options

- 100GBASE-DR (2m to 200m), 100GBASE-FR1 (2m to 2km) and 100GBASE-LR1 (2m to 10km)
- 400GBASE-FR4 (2m to 2km) and 400GBASE-LR4-6 (2m to 6km), both based on CWDM

802.3db: Establishes the multimode short reach optical specifications and options

- 100GBASE-VR1, 200GBASE-VR2 and 400GBASE-VR4 over 1, 2, or 4 pairs of multimode fiber up to at least 30m on OM3, 50m on OM4 and 50m on OM5
- 100GBASE-SR1, 200GBASE-SR2 and 400GBASE-SR4 over 1, 2, or 4 pairs of multimode fiber up to at least 60m on OM3, 100m on OM4 and 100m on OM5



802.3ck: Establishes the 100Gbps Lane Rate Electrical Specifications

Based on PAM4 (pulse amplitude modulation, 4 level) signaling

Defines solutions for 100Gbps, 200Gbps (2 x100G), and 400Gbps (4x100G) signaling interfaces

Leverages Reed-Solomon Forward Error Correction (RS-FEC) for all links Electrical Specifications:

Link Type	Standard	# of Lanes / Date Rate
Chip To Chip	100GAUI-1 C2C, 200GAUI-2 C2C, 400GAUI-4 C2C	1x100Gbps, 2x100Gbps, 4x100Gbps
Chip to Module	100GAUI-1 C2M, 200GAUI-2 C2M, 400GAUI-4 C2M	1x100Gbps, 2x100Gbps, 4x100Gbps
Copper Cable	100GBASE-CR1, 200GBASE-CR2, 400GBASE-CR4	1x100Gbps, 2x100Gbps, 4x100Gbps
Backplane	100GBASE-KR1, 200GBASE-KR2, 400GBASE-KR4	1x100Gbps, 2x100Gbps, 4x100Gbps





Chip to Chip and Chip to Module Both C2C and C2M define 100Gbps, 200Gbps, & 400Gbps

Chip to Chip (C2C):

ethernet alliance

• Enables approx. 25cm of reach with PCB traces (based on recommended 20dB Max channel loss)

Chip to Module (C2M):

- Enables the classic pluggable optical transceiver architecture
- Channel loss of 16dB, Host loss up to 11.9dB



NOTE—The number of lanes n is equal to 1 for 100GAUI-1, 2 for 200GAUI-2, and 4 for 400GAUI-4.

Chip to chip application



PCB based chip to module



Hybrid mix of PCB and cabled host chip to module





Copper Cable Implementation Application



100GBASE-CR1, 200GBASE-CR2, or 400GBASE-CR4 Link

Copper Cable Test Application (only module loss and host loss changes for C2M)



Cable assembly, host and test fixture insertion loss at 26.56Ghz



Copper Cable

Passive copper cables are a critical industry building block providing low power, low latency and low cost

Often used for "switch to server" links as well as "aggregation switch" links, i.e., high volume applications

Formfactors:

- SFP (1 channel): Enables 100GBASE-CR1
 SFP-DD and DSFP (2 channels): Enables up to 200GBASE-CR2
 QSFP (4 channels): Enables up to 400GBASE-CR4
 QSFP-DD* & OSFP* (8 channels): Enables up to 2x 400GBASE-CR4
- Cable assembly loss is 19.75dB at 26.5Ghz enabling a cable reach of 2 meters. This uses the same test fixtures as Chip to Module

Host loss up to 6.875dB

Second generation silicon is enabling longer cable reaches

*800GBASE-CR8 is in the process of being developed in IEEE P802.3df project





Example Module Compliance Board





Copper Cable 1x, 2x, 4x 100Gbps

100GBASE-CR1 Media Dependent Interface (MDI):

- SFP Formfactor: a single port (or PMD)
- SFP-DD or DSFP Formfactors: two ports (or PMDs)
- QSFP Formfactor: four ports (or PMDs)
- QSFP-DD or OSFP Formfactors: Eight ports (or PMDs)

200GBASE-CR2 Media Dependent Interface (MDI):

- SFP-DD or DSFP Formfactor: a single port (or PMD)
- QSFP Formfactor: two ports (or PMDs)
- QSFP-DD or OSFP Formfactors: four ports (or PMDs)

400GBASE-CR4 Media Dependent Interface (MDI):

- QSFP Formfactor: a single port (or PMD)
- QSFP-DD or OSFP Formfactors: two ports (or PMDs)

Breakout cable assembly examples:







Backplane

- Defines 100GBASE-KR1, 200GBASE-KR2, and 400GBASE-KR4 Physical Medium Dependent interface (PMD)
- Max channel loss is 28dB at 26.5Ghz, "ball to ball"



ethernet alliance



Orthogonal Backplane

Cabled Orthogonal Midplane

Cabled Backplane









Summary

- IEEE specifications are published to enable electrical and optical interfaces over a wide range of interfaces / media
- Allows optimization of implementations according to power, reach, etc. market demands, while leveraging the IEEE benefit of "plug and play" interoperability
- 100Gbps lane rates are challenging and careful attention to design detail is more important that it has ever been



Ethernet in Design (100Gb/s Lane Rate) Hardware Measurements

Adithya Muralidharan Staff Applications Engineer | Programmable Solutions Group Intel June 7, 2023



ethernet alliance

www.ethernetalliance.org

intel

Introduction

- Demonstration results of Ethernet variants at 53 GBd i.e., 100 Gb/s per lane showing
 - Bit Error Rate (BER) performance
 - Forward Error Correction (FEC) statistics
 - Ethernet Packet and Throughput Tests (400GE and 800GE)
- Across different media types such as
 - Backplane/Channel Evaluation Board
 - Direct Attach Copper (DAC)
 - Optical fiber
- Including different form factors
 - Quad Small Form-factor Pluggable Double Density (QSFP-DD)
 - Octal Small Form-factor Pluggable (OSFP)



Focus Areas

intel

- 802.3 ck Electrical Measurements
- 400GBASE-CR4 Hardware Measurements
- 800GAUI-8 Hardware Measurements



802.3 ck Electrical Measurements – Test Setup

- Intel Agilex F-Tile FHT Transceiver
 - 106.25 Gb/s (PAM4) Line Rate
 - QPRBS31 Pattern
- Channel Evaluation Board
 - Trace lengths ranging from 1.2 to 10 inches (up to 28 dB end-to-end loss)
- Connector Form Factor
 - OSFP

Agilex F-TILE Dev Kit with Channel Evaluation Board





intel

802.3 ck Electrical Measurements – Channel Model



- Agilex F-Tile FHT Transceiver
 - TX: 2.3778 dB
 - RX: 2.2324 dB
- OSFP to SMA Adapter
 - 1.9499 dB
- SMA Cable
 - 2.2832 dB

Nyquist @ 26.5625 GHz



intel

802.3 ck Electrical Measurements – Hardware Results

Trace Length (in.)	End-to-End Insertion Loss (dB) @ 106.25 GHz (Nyquist Rate: 26.5625 GHz)	Pre-FEC BER	Post-FEC BER	
1.2	15.3264	6.73E-12	0	
2.4	17.0664	4.52E-12	0	
4.8	21.1464	1.77E-11	0	
5.9	23.0464	5.30E-10	0	
7.1	25.0764	7.75E-09	0	
8.3	26.3964	5.72E-08	0	
9.5	27.2964	3.57E-07	4.25E-12	

Note : BER values recorded 10 minutes after starting transmission.

Data collected with Early Sample (ES) Silicon. Full characterization pending.

Things to Monitor

- As Insertion Loss increases, bit error rate increases
- There is an insertion loss point where FEC correctable errors starts to kick in





400GBASE-CR4 – Test Setup



Board 1

Board 2

- FPGA Board to Board Setup with 1-meter QSFP-DD DAC Cable
- 400G Ethernet consisting of four lanes with each lane running at 106.25 Gbps
- Reed Solomon (544, 514) Forward Error Correction Scheme





RX CDR Locked

400GBASE-CR4- Hardware Results (1) Board 2 – PHY Stats

Board 1 – PHY Stats

a construction and these internation					
PMA type: FHT	Ethernet mode: 400G-4	Client interface:	IAC Segmented	AN/LT: Yes	System PLL frequency: 870.0
FEC mode: RS(544,514) (CL134)	Enable CWBin: No	ADME:	res	PTP: Disabled	PMA reference frequency: 156.2
General Commands					
Enable Internal Loopback	Assert Soft Global	Reset	Assert Soft T	Reset	Assert Soft RX Reset
Vising Quartus Generated Example Desig	n				
Status Statistics Counters Testing Fe	atures				
Stop Reading All Status					
Stop Reading All Status					
PCS Status			* PHY Status		
Stop Reading All Status			PHY Status Enable PHY	n Read All Status	
Stop Reading All Status	Alignment Marker Loci	k •	PHY Status Enable PHY TX CLK kHz:	in Read All Status 435020	RX CLK IHz: 435000
Stop Reading All Status PCS Status C Enable PCS in Read All Status RX PCS Ready: TX Lanes Stable:	Alignment Marker Loci Frame Error:	k: •	PHY Status Enable PHY TX CLK kHz:	in Read All Status 435020	RX CLK IHz: 435000
Stop Reading AI Status PCS Status PC Status PC Enable PCS in Read AII Status RX PCS Ready. TX Lanes Stable: Deskewed status:	Alignment Marker Loci Frame Error:	k: • 0x0000000	PHY Status Enable PHY TX CLK kHz: PHY Lane 0	in Read All Status 435020	RX CLK IHz: 435000 PHY Lane 1
Stop Reading All Status PCS Status Enable PCS in Read All Status RX PCS Ready: TX Lanes Stable: Deskewed status:	Alignment Marker Loci Frame Error. Hi-BER:	k: • 0×00000000 0	PHY Status Enable PHY TX CLK kHz: PHY Lane 0 TX PLL Locked:	n Read All Status 435020	RX CLK IHz: 435000 PHY Lane 1 TX PLL Locked: ●
Stop Reading Al Status PCS Status Imable PCS in Read All Status RX PCS Ready: TX Lanes Stable: Deskewed status:	Alignment Marker Loci Frame Error: Hi-BER BER Count:	k: 0x0000000 0 0x0000000	PHY Status Finable PHY TX CLK kHz: PHY Lane 0 TX PLL Locked: RX CDR Locked:	n Read All Status 435020	RX CLK HHz: 435000 PHY Lane 1 TX PLL Locket: • RX CDR Locket: •
Stop Reading AI Status PCS Statuse Imable PCS in Read AI Status RX PCS Ready. TX Lanes Stable: Deskewed status:	Alignment Marker Loci Frame Error: Hi-BER BER Count: From Block Count	 k: 0×0000000 0 0×0000000 0×0000000 0×0000000 	PHY Status Finable PHY TX CLK kHz: PHY Lane 0 TX PLL Locked: RX CDR Locked: PHY Lane 2	n Read All Status 435020	RX CLK HHz: 435000 PHY Lane 1 TX PLL Locked: RX COR Locked: PHY Lane 3
Stop Reading AI Status PCS Status Reader CS in Read AI Status RX PCS Ready: TX Lanes Stable: Deskewed status:	Alignment Marker Loci Frame Error: Hi-BER BER Count: Error Block Count:	ic Ф Охооосоооо О Охооосоооо Охооосоооо Охоооооооо Охоооооооооо	▼ PHY Status ✓ Enable PHY TX CLK kHz: PHY Lane 0 TX PLL Locked: RX CDR Locked: PHY Lane 2 TX PLL Locked:	n Read All Status 435020	RX CLK Htz: 435000 PHY Lanc 1 TX PLL Locked: RX CDR Locked: PHY Lanc 3 TX PLL Locked:

* RS-FEC * Error Inser

> Total RS-FE Correcte

Things to Monitor:

- PHY Stats are in • good condition
- FEC uncorrectable • codewords is 0
- RS (544, 514) FEC ٠ can correct up to 15 symbol errors

FEC Tail

Data collected	with Early	Sample (E	S) Silicon.	Full characterization pending	



Board 1 –	-FEC Stats
-----------	------------

FEC				1
rror Insertion and Statistic	5			
tal RS-FEC Statistics				
Corrected Codewords A:	334025057	Corrected Codewords B:	334649542	
Uncorrected Codewords A:	0	Uncorrected Codewords B:	0	
Corrected Codewords Total	668674599	Uncorrected Codewords Total	0	
Corr. CWBin Cnt_0 A:	2224	Corr. CWBin Cnt_0 B:	1465	
Corr. CWBin Cnt_1 A:		Corr. CWBin Cnt_1 B:		
Corr. CWBin Cnt_2 A:		Corr. CWBin Cnt_2 B:		
Corr. CWBin Cnt_3 A:	8812	Corr. CWBin Cnt_3 B:		
Corr. CWBin Cnt_4 A:	9019	Corr. CWBin Cnt_4 B:	8784	
Corr. CWBin Cnt_5 A:		Corr. CWBin Cnt_5 B:	9377	
Corr. CWBin Cnt_6 A:		Corr. CWBin Cnt_6 B:		
Corr. CWBin Cnt_7 A:	150	Corr. CWBin Cnt_7 B:	137	
Corr. CWBin Cnt_8 A:	16	Corr. CWBin Cnt_8 B:	39	
Corr. CWBin Cnt_9 A:	2	Corr. CWBin Cnt_9 B:	8	Tail
Corr. CWBin Cnt_10 A:		Corr. CWBin Cnt_10 B:		i an
Corr. CWBin Cnt_11 A:	0	Corr. CWBin Cnt_11 B:		
Corr. CWBin Cnt_12 A:		Corr. CWBin Cnt_12 B:		
Corr. CWBin Cnt_13 A:		Corr. CWBin Cnt_13 B:		
Corr. CWBin Cnt_14 A:		Corr. CWBin Cnt_14 B:		
Corr. CWBin Cnt_15 A:		Corr. CWBin Cnt_15 B:		
Start Reading RS-FEC Sta	atistics	Reset RS-FEC Statistics		

Toolist T-16 (themnet heel POA Heed P (ed., f. bookt 1.0) PM-Agles (Seiner SOC De V Kit on localhost (USE-1)AOFB027R24C(PC3P0), @1[bs_nasterjatera_log_avalon_master_0 PM-Agles (Seiner SOC De V Kit on localhost (USE-1)AOFB027R24C(PC3P0), @1[bs_nasterjatera_log_avalon_master_0 Clean working directory on every process termination Delete working directory on tookit close IP informatior P Configuration and Other Infor PTP: Disabled FEC mode: RS(544.514) (CL134) Enable CABin: No ADME Yes PMA reference frequency: 156 2500 Enable Internal Loopback Assert Soft Global Reset Assert Soft TX Reset Assert Soft RX Reset Using Quartus Generated Example De Status Statistics Counters Testing Featur Stop Reading All Status * PCS Status Enable PCS in Read All Status Finable PHV in Read All Status RX PCS Ready: Alignment Marker Lock: TX CLK Hz: 435000 RX CLK kHz: TX Lanes Stable: PHY Lane 0 PHY Lane 1 Deskewed status: TX PLL Locked: TX PLL Locked RX CDR Locked RX CDR Locked PHY Lane 2 PHY Lane 3 TX PLL Locked: TX PLL Locked

RX CDR Lockert

Board 2 – FEC Stats

dut itag maste

Error Insertion and Statistic	8		
Total RS-FEC Statistics			
Corrected Codewords A:	300832784	Corrected Codewords B:	29903021
Uncorrected Codewords A:	0	Uncorrected Codewords B:	0
Corrected Codewords Total	599862994	Uncorrected Codewords Total:	0
Corr. CWBin Cnt_0 A:	1304	Corr. CV/Bin Cnt_0 B:	2
Corr. CWBin Cnt_1 A:		Corr. CWBin Cnt_1 B:	7
Corr. CWBin Cnt_2 A:		Corr. CV/Bin Cnt_2 B:	8
Corr. CV/Bin Cnt_3 A:	6806	Corr. CV/Bin Cnt_3 B:	9
Corr. CWBin Cnt_4 A:	8931	Corr. CV/Bin Cnt_4 B:	8
Corr. CWBin Cnt_5 A:		Corr. CV/Bin Cnt_5 B:	7
Corr. CWBin Cnt_6 A:	6#37	Corr. CWBin Cnt_6 B:	5
Corr. CWBin Cnt_7 A:	-59-	Corr. CWBin Cnt_7 B:	67
Corr. CWBin Cnt_8 A:	3	Corr. CWBin Cnt_8 B:	3
Corr. CV/Bin Cnt_9 A:		Corr. CV/Bin Cnt_9 B:	2
Corr. CV/Bin Cnt_10 A:	0	Corr. CWBin Cnt_10 B:	
Corr. CWBin Cnt_11 A:		Corr. CWBin Cnt_11 B:	
Corr. CWBin Cnt_12 A:		Corr. CWBin Cnt_12 B:	
Corr. CWBin Cnt_13 A:		Corr. CWBin Cnt_13 B:	
Corr. CWBin Cnt_14 A:		Corr. CWBin Cnt_14 B:	
Corr. CWBin Cnt_15 A:		Corr. CWBin Cnt_15 B:	
Start Reading RS-FEC Str	stistics	Reset RS-FEC Statistics	



400GBASE-CR4- Hardware Results (2)

Board 1 – MAC Stats

Board 2 – MAC Stats

Statistics Counters Names	TX Statistics	RX Statis		Statistics Counters Names	TX Statistics	RX Statistics
Frames < 64 bytes with CRC error	0	0		Frames < 64 bytes with CRC error	0	0
Oversized frames with CRC error	0	0		Oversized frames with CRC error	0	0
Packets with FCS errors	0	0		Packets with FCS errors	0	0
Frames >= 64 bytes with CRC error	0	0		Frames >= 64 bytes with CRC error	0	0
Multicast data frames with CRC error	0	0		Multicast data frames with CRC error	0	0
Broadcast data frames with CRC error	0	0		Broadcast date frames with CRC error	0	0
Unicast data frames with CRC error	0	0		Unicast data frames with CRC error		0
Multicast control frames with CRC error	0	0		Multicast control frames with CRC error	0	0
Broadcast control frames with CRC error	0			Broadcast control frames with CRC error	0	
Unicast control frames with CRC error	0	0		Unicast control frames with CRC error	0	0
Pause frames with CRC error	0	0		Pause frames with CRC error	0	0
64 Byte Frames (includes CRC field)	0	0		64 Byte Frames (includes CRC field)	0	0
65 - 127 Byte Frames	721186581	9612750224	Things to Monitor	65 - 127 Byte Frames	9612750224	721186581
128 - 255 Byte Frames	1201977635	0		128 - 255 Byte Frames	0	1201977635
256 - 511 Byte Frames	480791054	0	 Pandom Lonath 	256 - 511 Byte Frames	0	480791054
512 - 1023 Byte Frames	240395527	0	 Kunuum Lengin 	512 - 1023 Byte Frames	0	240395527
1024 - 1518 Byte Frames	1201977635	0	Packate Tastina	1024 - 1518 Byte Frames	0	1201977635
1519 - MAX Size Frames	0	0	ruckets testing	1519 - MAX Size Frames	0	0
Oversized Frames (SMAX Size)	0	0	 TV stats on board 	Oversized Frames (> MAX Size)	0	0
Multicast data frames without error	3846338433	9612750224		Multicast data frames without error	9612750224	3846328432
Broadcast data frames without error	0	0	1 matches with	Broadcast data frames without error	0	0
Upicast data frames without error	0	0	I MAICHES WITH	Unicast data frames without error	0	0
Multicast control from a without error	0	0	DV state an locard	Multicast control frames without error	0	0
Breedeast control frames without error	0	0	RX SIGIS ON DOGIG	Broadcast control frames without error	0	0
Broadcast control frames without error	0	0	0	Unicast control frames without error	0	0
Onicast control frames without error	0	0	L Z	Pause frames without error	0	0
Pause frames without error	0	0		Runt Packets (undersized)	0	0
Runt Packets (undersized)	0	0	 Similarly, IX stats 	Number of Frame Starts	9612750224	3846328432
Number of Frame Starts	3846328432	9612750224		Number of Length error	0	0
Number of Length error	0	0	on board 2	PFC frames with CRC error	0	0
PFC frames with CRC error	0	0		PFC frames without CRC error	0	0
PFC frames without CRC error	0	0	matches with RX	Payload data and padding octets	519088512096	2039275255541
Payload data and padding octets	2039275255541	519088512096		Frame Octets	692118016128	2108509167317
Frame Octets	2108509167317	692118016128	stats on board 1	Malformed packets	0	0
Malformed packets	0	0		Packets dropped due to error	0	0
Packets dropped due to error	0	0		Bad L/T field	0	0
Bad L/T field	0	0		4	III	

intel

800GAUI-8 Test Setup



- FPGA Dev Kit connected with Ethernet Protocol Tester using :
 - 800G-DR8 optical module and 3-meter optical fiber
- 800G Ethernet consisting of eight lanes from 2X FTILEs with each lane running at 106.25 Gbps

intel

800GAUI-8 Hardware Results (1)

FPGA Stats



Things to Monitor

- PHY/PCS Stats are in good condition
- FEC uncorrectable codewords is 0
- 800G Traffic running at
 - 100% Throughput on both TX and RX
 - 1518 Byte packet size

Traffic Results



ethernet alliance



800GAUI-8 Hardware Results (2)

Tester Stats

t	Empty Filte	r				C2 Re	fresh 🎦 Clear	Actio	ns
Ma	ain Aggregatio	on Port Conf	ig Resource Manage	r PFC Measu	rement L	1 Configuration BFD M	Nicro-BFD BGP	Qbv Stream C	onfiguration
	Port Name	Tags	Append Location to Port Name	Location	Online	Media Type	Link Status	Licensed Speeds	Performance Level
Þ	Port //1/1	Click to ad		//172.14.10	2	Ethernet 800 Gig Fiber	🥥 Up	100G/200	SWO-AHI-POR

Things to Monitor

- Tester showing 800GE link up status
- Tester running 800GE traffic at 100% throughput

Tr	raffic Aggregate View:Results 1 4 ×											
Port Traffic and Counters > Basic Traffic Results Change Result View - 🎦 🐘 🗛 - 🕪 🍕 🛛 1 of 1 🕨 📦											Port Traffic and Counters > Aggregate Port L1 Tx Ra	
Basic Counters Errors Triggers Protocols Undersize/Oversize/Jumbo PFC Counters User Defined Advanced Sequencing FEC Counters										Change Result View 🗸 🖓 🔢 🕨 📋		
	Port Name	Total Tx Count (Frames)	Total Rx Count (Frames)	Total Tx Count (bits)	Total Rx Count (bits)	Total Tx Rate (bps)	Total Rx Rate (bps)	Tx L1 Count (bits)	Rx L1 Count (bits)	Tx L1 Rate (bps)	Rx L1 Rate (bps)	Aggregate Port L1 Tx Rate
3	Port //1/1	17,097,945,183	15,336,663,985	207,637,446,302,352	184,285,354,443,760	789,596,877,384	789,486,803,080	210,373,117,531,632	186,739,220,681,360	799,999,998,304	799,999,277,153	
												240 480 240 560 160 640 80 720 0 Gbps 80
Σ		17,097,945,183	15,336,663,985	207,637,446,302,352	184,285,354,443,760			210,373,117,531,632	186,739,220,681,360			800.00

Traffic Aggregate View:Results 1 Traffic Aggregate View:Results 2 Validation Errors Log - 199 messages





Notices & Disclaimers

- Performance varies by use, configuration, and other factors. Learn more at <u>www.Intel.com/PerformanceIndex</u>.
- Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. No product or component can be absolutely secure.
- Your costs and results may vary.
- Intel technologies may require enabled hardware, software or service activation.
- © Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.



Ethernet in Design (100Gb/s Lane Rate) Clause 162, system validation challenges at TP2

John Calvin | June, 7 2023



ethernet alliance

www.ethernetalliance.org

TP2 Loss Related Challenges

Draft Amendment to IEEE Std 802.3-2022 IEEE P802.3ck 100 Gb/s, 200 Gb/s, and 400 Gb/s Electrical Interfaces Task Force IEEE Draft P802.3ck/D3.3 10 June 2022

Table 162–11—Summary of transmitter specifications at TP2 (continued)

Parameter	Subclause reference	Value	Units
Output jitter (max) J _{RMS} J3u ₀₃ J3u Even-odd jitter, pk-pk	162.9.4.7 162.9.4.7 162.9.4.7 162.9.4.7	0.023 0.115 0.125 0.025	ม บา บา บา

The recommended maximum insertion loss from TP0 to TP2 or from TP3 to TP5 including the test fixture is determined using Equation (162A–3), and illustrated in Figure 162A–2. The recommended maximum insertion loss allocation for the transmitter or receiver differential controlled impedance PCBs is consistent with the insertion loss from TP0 to TP2 or TP3 to TP5 and an assumed mated connector insertion loss of 1.6 dB. Note that the recommended maximum insertion loss from TP0 to TP2 or TP3 to TP5 and an assumed mated connector insertion loss of 1.6 dB. Note that the recommended maximum insertion loss from TP0 to TP2 or from TP3 to TP5 is 10.975 dB at 26.56 GHz.



and connector footprint vias.

Recommended TP0-TP2 is 6.875dB (host) + 1.6dB+2.5dB (mated connector + HCB) = 10.975dBSome larger packages (45mm) TP0-TP2 losses are more inline with 4dB (Package) + 6.875 (host) + 5.8dB (mated connector + HCB) = 16.72dB

The standards process recognized the challenges with TP2 and incorporated a couple useful changes.

- J3u was increased 10mUI to 125mUI to allow for higher loss packages and general challenges with performing 12Edge jitter operations after this much loss.
- J3u03 was added at the old J3u spec value of 115mUI. J3u03 limits the J12Edge Jitter operations to full swing (3 level transitions) which are less susceptible to SNR issues after a high loss channel.

TP2 Test Fixture loss

Draft Amendment to IEEE Std 802.3-2022 IEEE P802.3ck 100 Gb/s, 200 Gb/s, and 400 Gb/s Electrical Interfaces Task Force

IEEE Draft P802.3ck/D3.3 10 June 2022

(162B-4)

162B.4 Mated test fixtures

The TP2 or TP3 test fixture and the cable assembly test fixture are specified in a mated state illustrated in Figure 92–18. The mated test fixtures specifications are given below.

162B.4.1 Mated test fixtures insertion loss

The insertion loss of the mated test fixtures shall meet the values determined using Equation (162B–3) and Equation (162B–4).

$$ILdd(f) \le ILdd_{MTFmax}(f) = \left\{ \begin{array}{cc} 1.185 \sqrt{f} - 0.072f + 0.007f^2 & 0.01 \le f < 40\\ 1.915f - 60.78 & 40 \le f \le 50 \end{array} \right\}$$
(162B-3)

 $ILdd(f) \ge ILdd_{MTFmin}(f) = 1.0225(0.0656\sqrt{f} + 0.164f)$

for 0.01 GHz $\leq f \leq$ 50 GHz

ethernet alliance





Figure 162B-2 illustrates allowable test fixture SDD21 tolerances

Referice SDD21 is -6.6dB, Min is -4.8dB Max is -9.13dB Top SDD21 illustrates an example of a conformant test fixture

P802.3ck Electrical/Jitter Specs in Cabled systems (CR)



Figure 162–2—100GBASE-CR1, 200GBASE-CR2 or 400GBASE-CR4 link

Draft Amendment to IEEE Std 802.3-2022 IEEE P802.3ck 100 Gb/s, 200 Gb/s, and 400 Gb/s Electrical Interfaces Task Force

162.9.4.7 Output jitter

Output jitter is characterized by four parameters: J_{RMS_1} J3u, J3u₀₃, and even-odd jitter. These parameters are calculated from measurements with a single transmit equalizer setting to compensate for the loss of the transmitter package and host channel. The equalizer setting is chosen to minimize any or all of the jitter parameters.



TP2 is performed with TX FIR only and a 40GHz 4'th order Bessel Thomson instrument response.

Large levels of data-dependant jitter (DDJ) are present after traversal through these loss elements and need to be compensated with any configuration of the Tx silicon's TX-FFE to achieve the lowest possible 12Edge jitter values.

No Reference Receiver equalization such as CTLE or DFE is permitted for TP2. Only TX Fir optimization is allowed.



12Edge Operations Function on closed eye's

Run Stop Single 256 GSa/s 256 kpts Bandwidth	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	Auto	
Jitter D PAM4 Jitter D Results (Measure All Edges) D				(1)
f1:4Port(m1) All tables All graphs				
f1 J3u Table (832 Transitions) □ f1 Jrms Table (832 Transitions)	f1 L0 to L1	f1 L1 to L0	f1 L2 to L0	f1 L3 to L0
Fron To L0 To L1 Fron To L0 To L1 All 111.771 mUI All 18.7895 mUI	130:015:035 13:0015:035 13:001 17:015:001 10:01 10:01	1 alisuolis. 853 133 133 13 14 14 14 1 14 1 1 1 1 1 1	• J3u: 98.4109 mUI • J4u: 98.4109 mUI	J3u: 84.2474 mUI J4u: 84.2474 mUI
L3 84.2474 mUI 94.6538 mUI 115 L3 17.2141 mUI 17.6347 mUI 20.	10.951 min 11.11		• Jrms: 16 6 380 m01	• Jrms: 172141 (101 m) (11 m)
f1 J4u Table (832 Transitions) f1 EOJ Table (832 Transitions)	1 L0 to L2	f1 L1 to L2	f1 L2 to L1	f1 L3 to L1
All 123.447 mUl	33u: 100.749 mUI J4u: 100.749 mUI	• 13u: 110.844 mUI • 14u: 110.844 mUI • 16u: 110.844 mUI	D3u: 113.575 mUt J4u: 113.575 mUt J6u: 113.575 mUt	• J3u: 94.6538 mUI • J4u: 94.6538 mUI • J6u: 94.6538 mUI
L3 84.2474 mUI 94.6538 mUI 115	16 V Dims: 17.6980 111 11 1		• Jrms: 20.3653 mu	• Jrms: 17 63 5 10 11 11 11
From To L0 To L1	Transitions: 832	Transitions: 832	Transitions: 833	Transitions: 833
Ali 123.447 mUl	J4u: 88.5119 mUI J4u: 88.5119 mUI J6u: 88.5119 mUI	J30: 86.2435 DU1 J40: 86.2435 DU1 J60: 86.2455 DU1	■ J3u: 112.654 mUI ■ J4u: 112.654 mUI ■ J6u: 112.654 mUI	 J3u: 115.287 mU J4u: 115.287 mU J6u: 115.287 mU
L3 84.2474 mUI 94.6538 mUI 115		The second se		
Image: https://www.end/action.com/action/	$-114 \mu V \checkmark 0 \land \oplus \ll \varphi$	J		
				456 mV
Keal-Time Eye				23
				-0.0 V
e a s		-		-456 mV
-18.8 ps -15.1 ps -11.3 ps	-7.53 ps -3.77 ps 0	0 s 3.77 ps 7.53 g	ps 11.3 ps 15.1 p	18.8 ps 💷
Real-Time Eye				
6.82082 MUI				-0.0 X
				-0.0 V
-18.8 ps -15.1 ps -11.3 ps	-7.53 ps -3.77 ps 0	0 s 3.77 ps 7.53 p	os 11.3 p <u>s</u> 15.1 ps	-456 mV 18.8 ps f1

Ethernet Alliance: Clause 162 validation challenges at CR's TP2

ethernet alliance

TP2 measurement use-case is important



From the standpoint of being able to confirm electrical interoperability of a CR Transmitter port by inserting an HCB and evaluating the signal properties, TP2 is important for system validation

The Jitter values at TP2 are currently based on a conservative set of loss values.

Under most conditions 10.975dB-15dB net loss, the current J3u value of 125mUI **is** attainable with a 5 tap FIR. Beyond 15dB requires careful study of advanced TX FIR capability, as margin disappears quickly.

Total Path Loss (dB)	TX FIR	J3U (mUI)	JRMS (mUI)
8	0/ 0.03/ -0.15/ 0.7/-0.08	75	10
15	-0.01/0.05/ -0.18/ 0.55/ -0.21	120	16

Figure 162B–3—Mated test fixtures differential-mode to differential-mode insertion loss

ethernet alliance

Even-Odd-Jitter (EOJ) Measurements

COMBINATION OF FBAUD, PATTERN LENGTH, CRU LOOP BW

In practice, the EOJ result is impacted by a physical CDR loop BW and the pattern harmonics

- 1. EOJ from DUT: CDR responds to the EOJ in the signal, is filtered by the Jitter Transfer Function (JTF) low-pass response, and moves the phase of the clock.
 - o Spectral Component = Fbaud / (2*PatternLength) = 53.125 GBd/16,382 symbols = <u>3.24 MHz (in-band)</u>
 - o Normal distribution, increases/decreases the amplitude of the EOJ component based on phase relationsl
 - o But EOJ algorithm always records the <u>maximum</u> of the of the 12 edges (worst case result), so EOJ result wi always be increased due to the component (sub-harmonic) of EOJ that falls within the loop BW of the CDR.
 - The reference clock recovery unit (CRU) used in the measurement acts as a highpass jitter filter with a corner frequency of 4 MHz and a slope of 20 dB/decade



CDR Src: DIA Symbol Rate: 53.125000	GBd		6	
JSA Jitter Spectrun	n Analysis Results			
Jitter Spectrum: Emulated Spectral Measurement Band Limits: 1.500	kHz to 26.25 MHz	Avg: 4/16 E	dit	
Measurements	Spectrum Peaks			
* Results are bandlimited.	Frequency 🛆	Magnitude		
TJ (rms)*: 225.94 fs	613.67 kHz	5.89 fs		
DJ (rms)*: 211.05 fs	1.0078 MHz	1.98 fs		
RJ (rms)*: 80.68 fs	1.0857 MHz	740 as		
	1.6213 MHz	4.18 fs		
	3.2429 MHz	62.64 fs		
	3.8566 MHz	11.64 fs	T	
	4.2508 MHz	3.22 fs		
	4.8643 MHz	11.26 fs		
	5.4779 MHz	8.49 fs		
	5.8721 MHz	9.28 fs		
	6.4858 MHz	125.41 fs		

Sub-Harmonics of EOJ are created based on Fbaud and Pattern Length:

- o Baud Rate: 53.125 Gbd
- PRBS13Q Pattern Length: 8191
 Symbols
- o Spectral Component = Fbaud / (2*PatternLength) = <u>3.24 MHz</u> (in-band)



What can be done?

Since "Real" systems don't have repetitive patterns that create this type of issue, it is our belief that any impairment of EOJ due to this relationship is <u>unintended by the standard</u>.

The physical PLL/Harmonic "interaction" effect can be mitigated *if EOJ is measured with*:

- 1. A lower CR Loop BW.
 - Allow CR loop BW to be lowered from 4 MHz to x MHz (Example: ~ 100 kHz).=> Simple
- 2. A shorter pattern such as PRBS9Q
 - Fbaud / (2*PatternLength) = 53.125 GBd / (2*511) = 52 MHz >> 4 MHz CR loop BW
- 3. An instrument with an oversampled digital PLL such a Real-Time Oscilloscope

Final .3ck spec says..

- The test pattern is either PRBS13Q or alternatively PRBS9Q. PRBS9Q is defined in 120.5.11.2.a. Meeting the even-odd jitter requirement with only one pattern is sufficient.
- b) If the test pattern is PRBS13Q, the corner frequency of the clock recovery unit (CRU) is set to 4 MHz or to 1 MHz. Meeting the even-odd jitter requirement with only one CRU bandwidth is sufficient.

Instrument Correlation

CR-Measurement (TP2	UXR(11.50)	DCA (7.50)	Difference	Comments
JRMS(mUI)	24.8	23	1.8	Equivalent-time instruments have added advantage of 1.5 ENOB better sensiti
J3u(mUI)	170.8	161	9.8	Real-time instruments will always report slightly higher Jnu over a Equivalent-
EOJ(mUI)	31.3	45	-13.7	The consecutive cycle to cycle nature of EOJ favors a Real-time instrument
RLM	0.961	0.969	-0.008	Should expect a good match with closed eye's favoring RT.
Vf(mV)	325	327	-2	How current is your calibration
SNDR(dB)	33.2	32.9	0.3	Within normal run -to-run variances, these should be similar

- Differences in the two instruments are to be expected within the confines of architectural and fidelity differences.
- In terms of which instrument is closest to the "correct value" that depends.

ethernet alliance



Summary

- The EA interoperability program is an ideal forum to evaluate both prototype and finished designs. Early exposure to the test program is highly beneficial for both Design Validation teams as well as the Test Instrumentation teams.
- Many of the 802.3ck early Systems tested have "interoperability opportunities".
- IEEE 802.3ck electrical validation has certain tight validation margins and a "one instrument does it all" approach is unlikely to serve the best overall results. It's ok to mix and match the instruments to get access to the breadth of best margin results.



Ethernet in Design (100Gb/s Lane Rate)

Pavel Zivny | June, 7 2023



ethernet alliance

www.ethernetalliance.org

HSN Plugfest Highlights (May 1-5, UNH-IOL)

- Alliance hosted a week-long High Speed Networking (HSN) plugfest at UNH
- 18 member companies participated in this event helping to improve ecosystem
 interoperability with a special focus on interoperability of Ethernet devices supporting
 data rates of 25-800 Gb/s
- Event was also **open to non-members**
- **Blog** highlighting technical details is coming soon
- More plugfests planned for **2023**, stay tuned for details





Physical layer tests considered at plugfest 1

Electrical standards: CR ✓, AUI C2M√ 802.3ck 100 Gb/s/lane √ 802.3ck 50 Gb/s/lane √

For optical standards, see next page



Physical layer tests for future consideration for Phy measurements

Optical standards

- 100 Gb/s/lane SM IMDD √/×
 - 50G interest? √/×
 - 200 G interest? √/×

100 Gb/s/lane MM IMDD (802.3db, others)

• 50 G interest √/×



Tektronix[®]

Physical layer test stations



Physical layer 802.3ck fixturing



Test & Measurement vendors supplied host and module fixtures for proving the standards

CMIS (Common Management Interface Specification) and older management were used to communicated to DUTs



Physical layer test: Station setup

Plugfest setup – Tektronix station:

Two stations of 802.3ck TX Test supporting setups, each based on Tek DPO70002SX oscilloscopes



Physical layer test: CR jitter example



Figure 162–2—100GBASE-CR1, 200GBASE-CR2, or 400GBASE-CR4 link

From IEEE 802.3 2022/12 specification



Physical layer test: compliance table example

Directly based on standards compliance characteristics, e.g.

Table 162–11—Summary of transmitter specifications at TP2

Parameter	Subclause reference	Value	Units
Signaling rate, each lane (range)		$53.125\pm50~\text{ppm}^{\text{a}}$	GBd
Differential pk-pk voltage with Tx disabled (max) ^b	93.8.1.3	30	mV
DC common-mode voltage (max) ^b	93.8.1.3	1.9	V
AC common-mode peak-to-peak voltage (max) Low-frequency, VCM _{LF} Full-band, VCM _{FB}	162.9.4.4	30 80	mV mV
Differential pk-pk voltage, v _{di} (max) ^b	93.8.1.3	1200	mV
Effective return loss, ERL (min)	162.9.4.8	7.3	dB
Common-mode to common-mode return loss, RLcc (min)	162.9.4.9	See Equation (162–6)	dB
Common-mode to differential-mode return loss, RLdc (min)	162.9.4.10	See Equation (162–7)	dB
Transmitter steady-state voltage, $v_f(\min)$ Transmitter steady-state voltage, $v_f(\max)$	162.9.4.1.2	0.387 0.6	V
Linear fit pulse peak ratio, R _{peak} (min)	162.9.4.1.2	0.397	
Level separation mismatch ratio R_{LM} (min)	162.9.4.2	0.95	
Transmitter output waveform absolute value of step size for all taps (min) absolute value of step size for all taps (max) value at minimum state for $c(-3)$ (max) value at maximum state for $c(-2)$ (min) value at minimum state for $c(-1)$ (max) value at minimum state for $c(0)$ (max)	$\begin{array}{c} 162.9.4.1.4\\ 162.9.4.1.4\\ 162.9.4.1.5\\ 162.9.4.1.5\\ 162.9.4.1.5\\ 162.9.4.1.5\\ 162.9.4.1.5\end{array}$	$\begin{array}{r} 0.005\\ 0.025\\ -0.06\\ 0.12\\ -0.34\\ 0.5\end{array}$	
value at minimum state for $c(1)$ (max)	162.9.4.1.5	-0.2	

Partial list only, see next section

Physical layer test: CR output jitter compliance

IEEE 802.3ck CR jitter specification:

Physical Medium Dependent (PMD) sublayer and baseband medium,

type 100GBASE-CR1, 200GBASE-CR2, and 400GBASE-CR4 are a good example of the jitter measurements in the standard.

162.9.4.7 Output jitter specification, Table 162–11—Summary of transmitter specifications at TP2

Output jitter (max)			
J _{RMS}	162.9.4.7	0.023	UI
J3u J3u	162.9.4.7	0.115	UI
Even-odd jitter, pk-pk	162.9.4.7	0.025	Ū

Where both the J3u and Even-odd jitter, pk-pk are different from previous versions of the standard.

Let's look at the differences...



J3U, J3U₀₃, J

Per IEEE 802.3, J3u is calculated using the measurement method for J4u (see 120D.3.1.8.1), except that J3u is defined as the time interval that includes all but 10–3 of fJ(t), from the 0.05th to the 99.95th percentile of fJ(t). J3u03 is calculated the same way as J3u except that the jitter calculation uses only transitions R03 and F30 in Table 162–13.

Why does the standard define both the J3u and the J3u03 as mandatory?

Even Odd jitter: Even-odd jitter EOJ is the maximum of the 12 measurements of even odd jitter.

EOJ is one of the few measurements that can be measured with the standard patter PRBS13Q, or with a shorter PRBS9Q pattern. Additionally for PRBS13Q the clock recovery PLL Loop bandwidth can be lowered to 1 MHz.

Why does the standard allow different patterns and even PLL loop bandwidth?

- The reason behind these changes is that caution needs to be exercised with the pattern length duration approaching the PLL loop bandwidth time constant.
- The precautions of the standard were reasonable: the J3u and J3u₀₃ measurements correlated reasonably well. The DUTs struggle with EOJ measurement in some cases, and additional (shorter) pattern is welcome there.



Future of the plugfest, feedback

- The 2023/H1 plugfest was aligned with an InfiniBand[™] plugfest. As Ethernet and IB / RoCE share similar physical layer a further future cooperation is being considered by both parties.
- The plugfest was well attended and generated good feedback. It is our intend to offer future plugfests, e.g. in 2024/H1, perhaps also in 2023/H2.





ethernet alliance

www.ethernetalliance.org

Thank you for watching!

If you have any questions or comments, please email admin@ethernetalliance.org



