IEEE 802.3ap "Backplane Ethernet" Overview

May 25, 2005
6 Gig SAS Meeting

Ali Ghiasi, Chief Arcitect
aghiasi@broadcom.com

History of Backplane Ethernet

- Ethernet previously did not standardize backplane interface as Ethernet traditionally focuses on the box interface.
  - Blade servers are changing the boundary of box / backplane interface
  - XAUI will be the closet standard in this space
- In Nov. 2003 IEEE 802.3 approved Backplane Ethernet Study group.
- In March 2004 Backplane Ethernet task force approved and 802.3ap officially started.
- Latest draft of 802.3ap is 0.9 and 0.91 will be available by June 6, located in member only area.
  - All the presentation and the task force material are available at http://www.ieee802.org/3/ap/index.html

BROADCOM. Connecting Everyone

BROADCOM. Connecting everything

What is Backplane Ethernet

- Leverage Existing Ethernet fabric and MAC
- Use existing Ethernet standard to develop an interface optimized for backplane
- Not a chassis or connector specification
- Will not define mechanical, thermal, or material
- Primary area of focus will be:
  - Electrical specifications
  - Channel model specifications

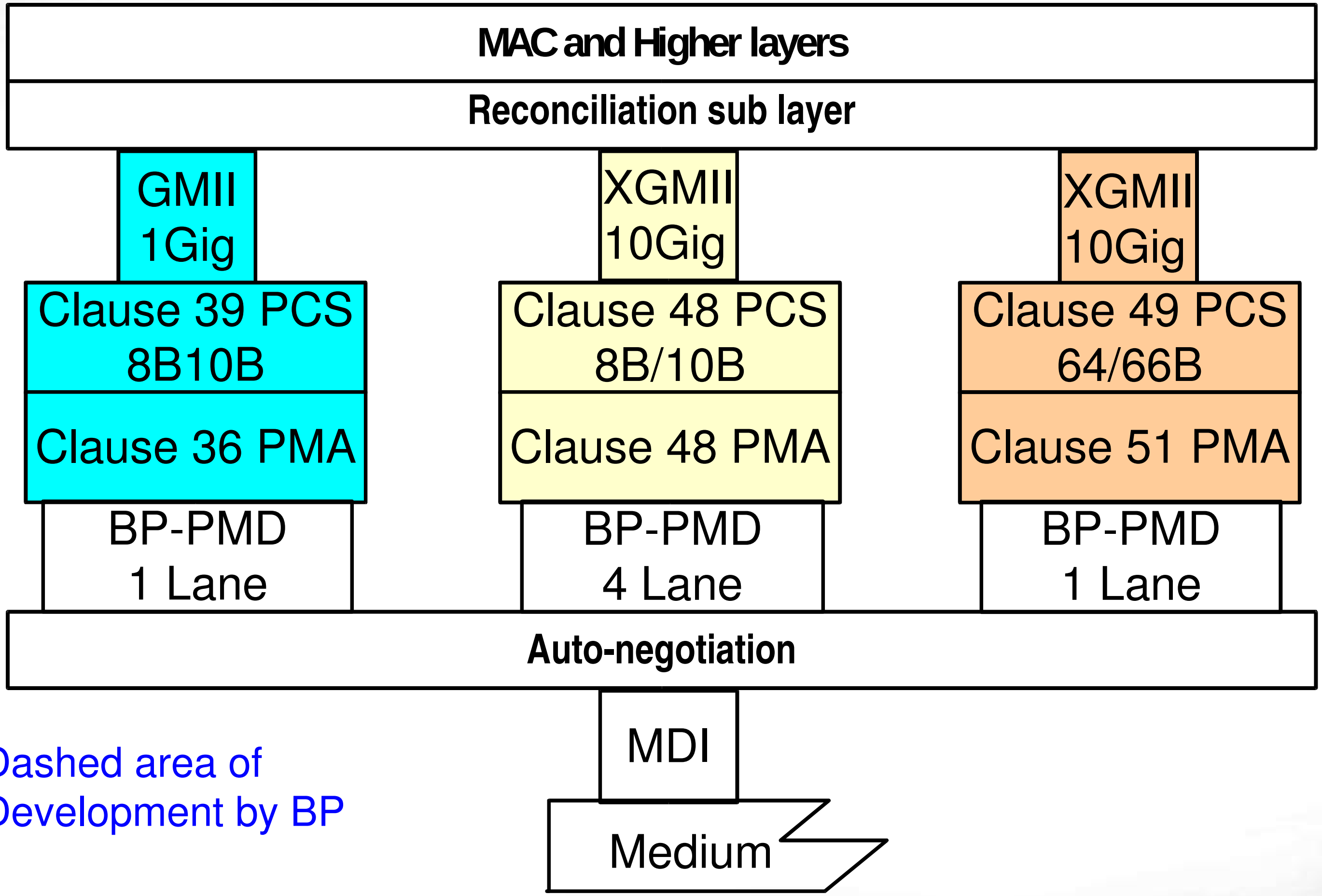
Justification for Backplane Ethernet

- The convergence of network and computing created Blade-center.
- In blade center environment about half the Ethernet are internal and operates over the backplane.
  – According to Gardner Report published IEEE Spectrum Apr. 2005 blade computing grows from $1.1B to $3.1B in 2009.
- Already significant amount of blade computing are shipping.
  – Many of these product use standard Ethernet chips and adopt them to operate over the backplane.
  – Majority of backplane Ethernet ports shipping with blade computing today is 1 Gig E but the migration to 10Gig E has started
    - At 10Gig commodity Ethernet chips may not operate reliably over the backplanes and proprietary solution will not have Ethernet economy of scale.

IEEE 802.3ap Objective as Specified in the PAR

- Preserve the 802.3 Ethernet frame format at the MAC client interface
- Preserve Min. and Max. frame size
- Support existing media independent interface
- Support operation over 1m of improved FR4 with 2 connectors
  - Define a 1 Gb/s PHY
  - Define a 10 Gb/s PHY
    - Single and 4 lane
- Automatic speed and feature negotiation.
- Support BER of 1E-12 or better
- Meet CISPR/FCC Class A.

Backplane Ethernet Layer Model



Dashed area of
Development by BP

1000Base-KX 1000Base-KX4 1000Base-KR

BROADCOM. Connecting Everyone

BROADCOM. Connecting everything

Auto-Negotiation

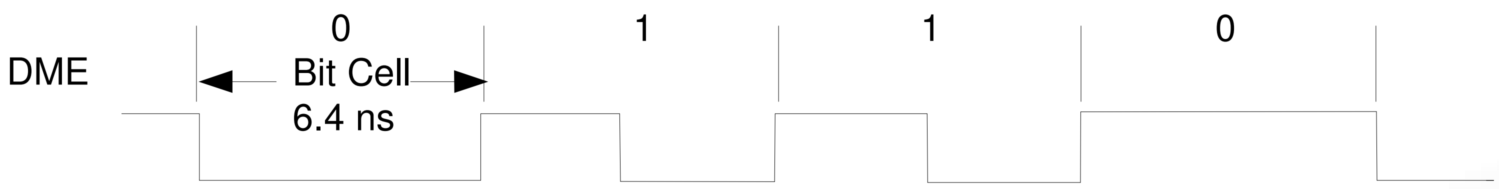
- IEEE 802.3ap defines scheme to Auto-Negotiate between KX, KX4, and KR.
- Auto-Negotiation is optional and host instead may use parallel detect
- Auto-Negotiation management definition is based on the IEEE Clause 28 " Definition of auto-negotiation for twisted-pair link Segment".
  - The actual signaling is not based on the Fast Link Pulses (FLPs or OOB), but instead is based on differential Manchester coding which is more suitable for high speed backplane.

XCVR Friendly Auto-neg

- FLP "OOB" not XCVR Friendly for > 3 Gb/s
  - Requires turning on/off transmitter very quickly ~100 ns pulse.
  - Receiver must have envelope detect to detect bursts.
  - PLL must acquire lock and adapt at the start of each training sequence
- Attributes of XCVR Friendly Auto-neg
  - Provide continuous clock information
  - DC balance
  - Baudrate low enough to operate on a untrained link
  - Higher that the signal minimum passband
  - Simple decode and encoder
  - Minimum additional requirement on the transceivers
- 6 Gig SAS can also benefit from a XCVR friendly Auto-Neg!

Auto-Neg Based on Differential Manchester Encoding (DME)

- **During Auto-neg DME operates at 312.5 Mbaud (6.4 ns ck-ck)**
  - **One symbol is equivalent to 33 symbols at 10.3125**
  - **Encoding rules bit cell always has transition**
  - **Generates DC-Balance signal at any one of the three rate currently defined**
  - **Bit cell center has transition for 1 but not for 0**
  - **Below example of DME and the equivalent SSP signal**



Ref. 3

Training Frame Structure

- **10GBASE-KR transmitter during training sends out DME codes at ¼ of baudrate of 10.3125 Gb/s.**
  - **Training frame structure is 100 octets in length as shown below:**

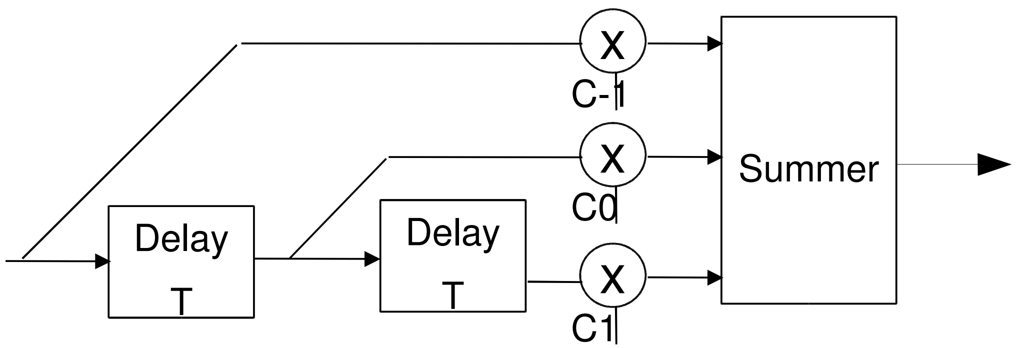
Octets		
4	Frame Marker	0xFFFF0000 @10.3125 Gbaud
16	TX FFE Cof. Update	DME Coded 6.4 ns transitions
16	Status Support	DME Coded 6.4 ns transitions
64	Training pattern	64 Byte pattern send @10.3125 Gbaud

Compliance Points

- Transmitter – Normative
  - Must meet electrical, jitter, and min tap granularity
- Channel – Informative
  - Unlike Cat-x cable widely available in the market and standardized by the TIA you can't buy an IEEE / TIA complaint backplane.
  - In contrast to CAT-x and Single Mode Fiber (SMF) backplanes have significantly greater degree of freedom with variation due to material, trace geometry, and connectors.
  - IEEE BP has elected to be instead very precise regarding SerDes specifications so the designer can build manufacturable backplane
- Receivers – Normative
  - Must meet worst case electrical input, jitter, and impulse response

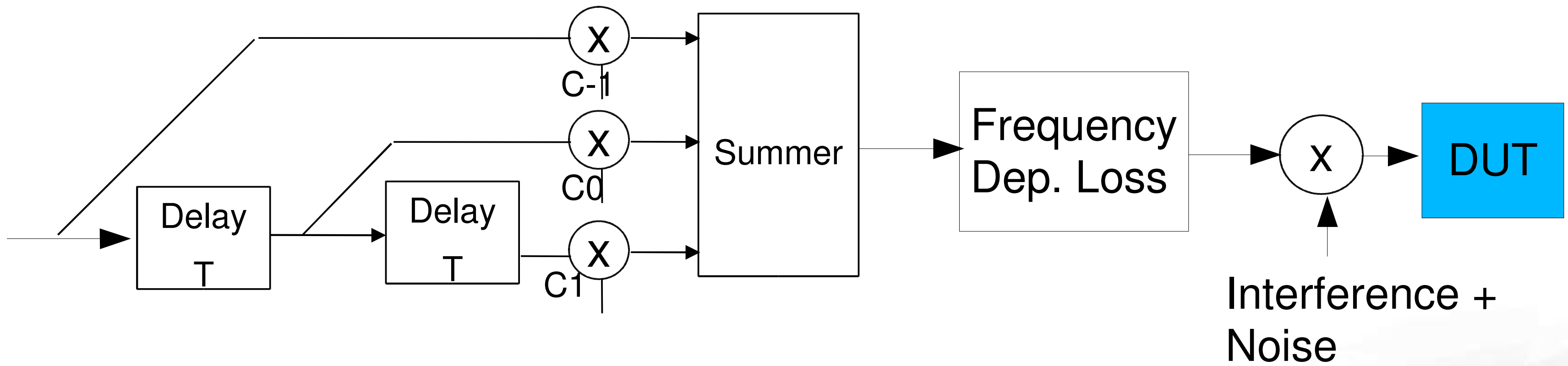
Transmitter Architecture

- **At minimum a three tap FFE will be implemented**
  - **C-1 8 steps from 0 to -7/40**
  - **C1 16 steps from 0 to -15/40**



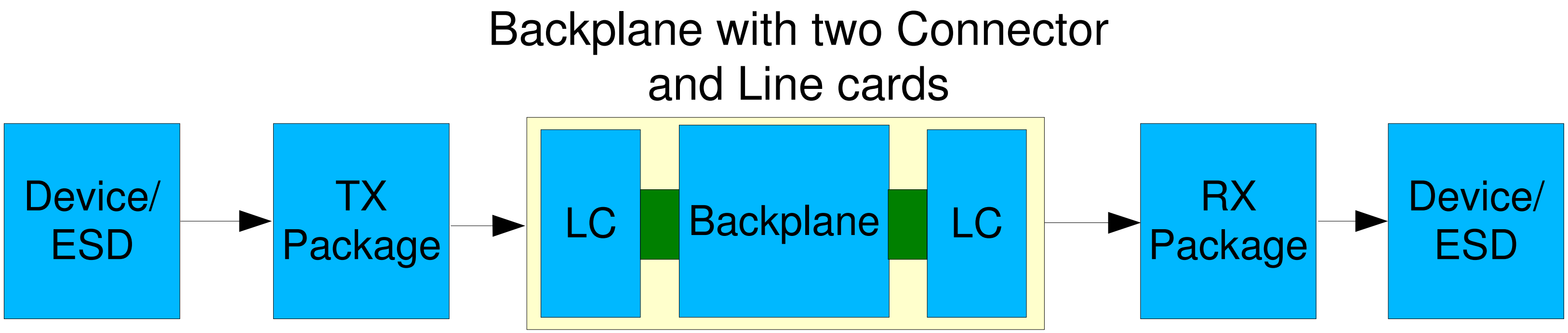
Receiver Compliance

- IEEE does not specifically defines an receiver, most common implementations are FFE/DFE based.
- Receiver will be tested with a complaint transmitter plus channel stressor.
  - Receiver compliance testing still is under refinement, but it looks like:



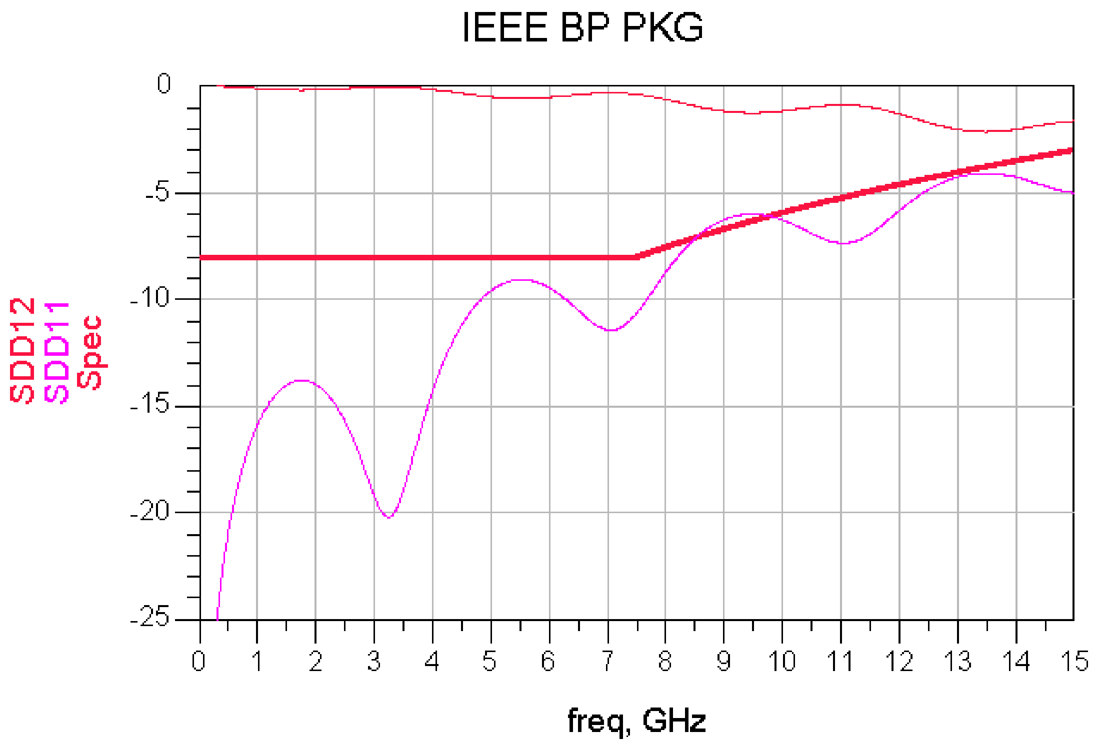
Channel Model

- **Through response is the cascaded s-parameters of each section**
  - **Aggregressors are added to the receive signal.**



Proposed Package Model

- **Capacitive package Model**
  - **Mask limits may change**



Eqn Spec=if(freq<7.5e9)then -8 else (-8 + 16.67*log10(freq/7.5e9))

Ref. 1

BP Transmitter Electrical Specifications

Parameter	*1000BASE-KX*	*10GBASE-KX4*	*10GBASE-KR*	*Units*
Signaling Speed, Per Lane	1.25 ± 100 ppm	3.125 ± 100 ppm	10.3125 ± 100 ppm	GBd
Differential Peak-Peak Output Voltage	800 to 1600	800 to 1600	120	mV
DC Common Mode Voltage	-0.4 to 1.2	-0.4 to 1.9	TBD	V
Differential Output Return Loss (Min)	10 [1]	10 [1]	8 [2]	dB
Output Jitter peak-peak (DJ)	0.1	0.17	0.15	UI p-p
Random Jitter (RJ)	NA	NA	0.15	UI p-p
Total Jitter	0.25	0.35	0.3	UI p-p

1. (RL >625 MHz) = 10 -10xLOG10(f/625), 625MHz <= f <= 2000 MHz

0. (RL >7.5 GHz) = 8 -16.6xLOG10(f/7.5), 7.5 Ghz <= f <= 15 GHz

Ref. 1

BP Receiver Electrical Specifications

Parameter	*1000BASE-KX*	*10GBASE-KX4*	*10GBASE-KR*	*Units*
Signaling Speed, Per Lane	1.25 ± 100 ppm	3.125 ± 100 ppm	10.3125 ± 100 ppm	GBd
BER	1.00E-012	1.00E-012	1.00E-012	
Differential Input Sensitivity	200	TBD	TBD	mV
Receiver Coupling	AC	AC	AC	
Differential Input Peak-Peak Amplitude (max)	1600	1600	1200	V
Jitter Tolerance [1]	baud/1667	baud/1667	baud/1667	
Differential Output Return Loss (Min)	10 [2]	10 [3]	8 [4]	dB
Common Mode Return Loss	6 [2]	NA	NA	dB

1. Jitter tolerance mask -20 dB/Dec from baud/25000 to baud/1667, then 0.1 UP up to 20 Mhz.

0. From 50MHz to 625 MHz

0. (RL >625 MHz) = 10 -10xLOG10(f/625), 625MHz <= f <= 2000 MHz

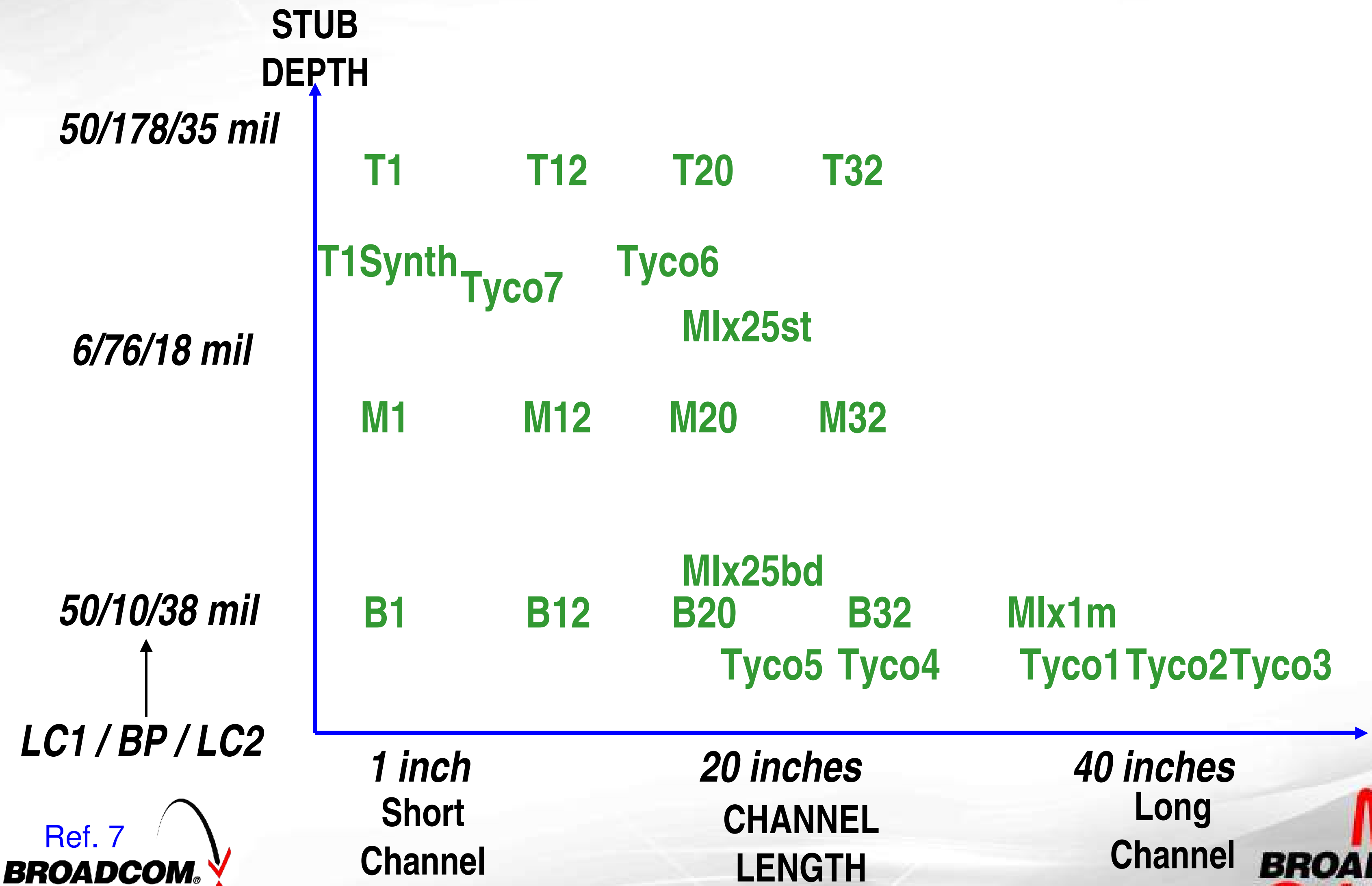
0. (RL >7.5 GHz) = 8 -16.6xLOG10(f/7.5), 7.5 Ghz <= f <= 15 GHz

Ref. 1

BROADCOM Connecting Everyone

BROADCOM Connecting everything

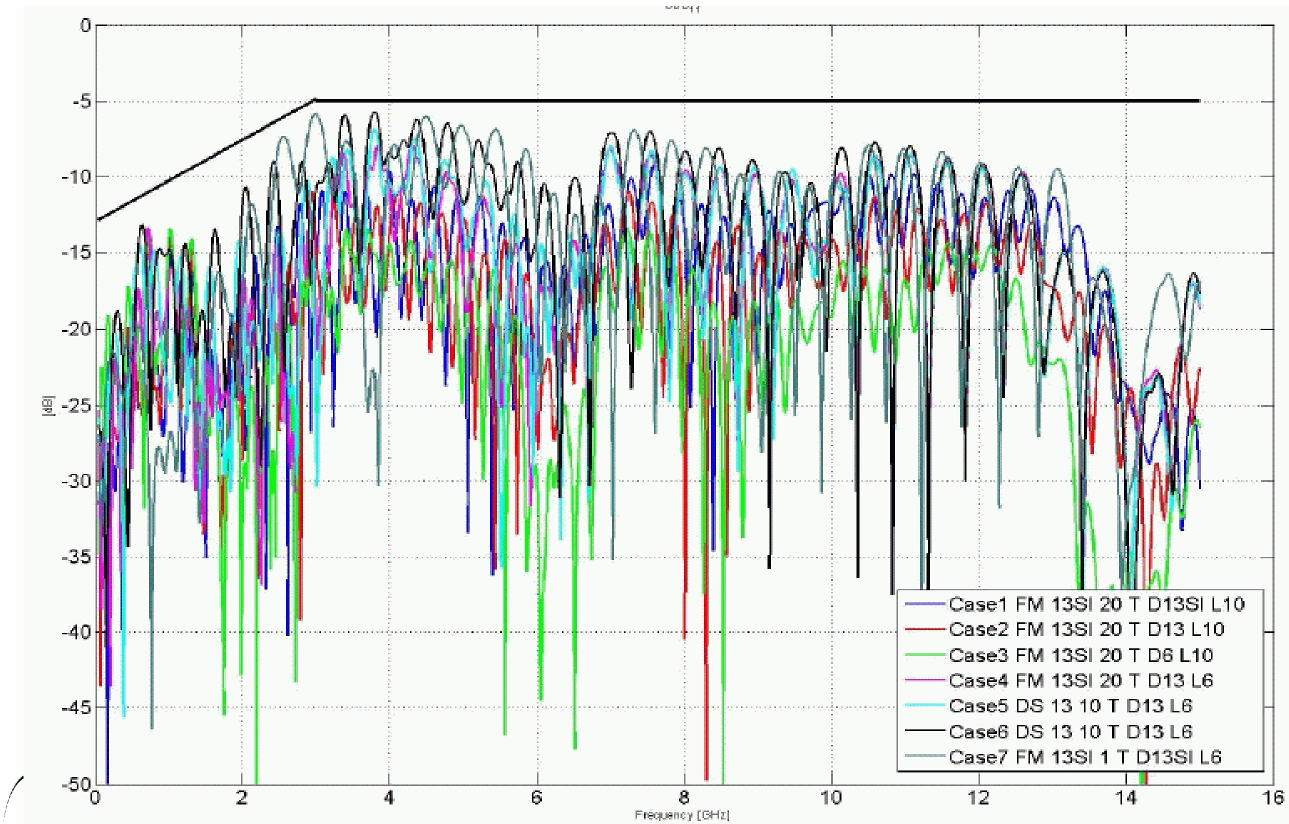
Overview of the IEEE Backplanes



Ref. 7

IEEE Tyco Channels

- **Return loss for several backplane including ATCA**
  - **Channel length 13-40” based on improved FR4**



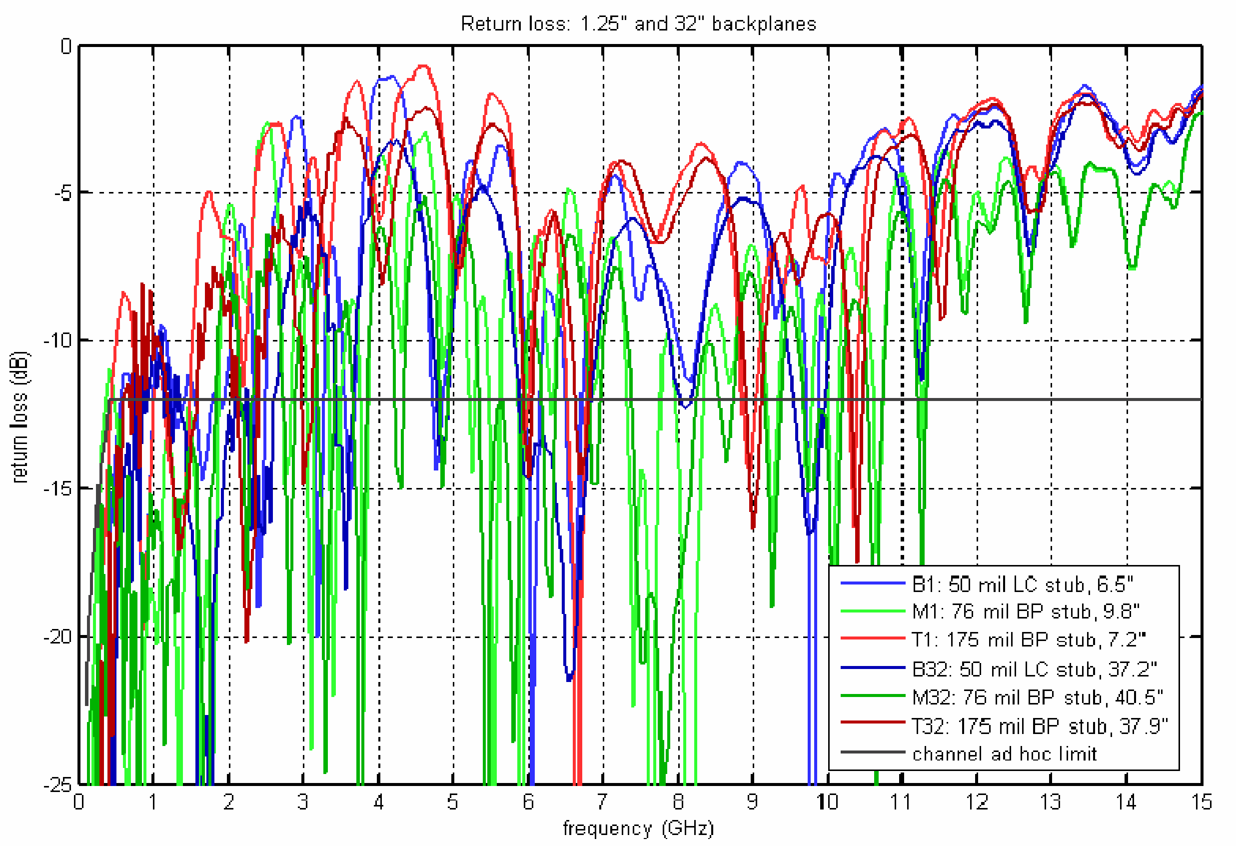
Ref. 6

BROADCOM. Connecting Everyone

BROADCOM. Connecting everything

IEEE Intel Channels

- **Return loss for several ATCA Backplanes**
  - **Channel length 1.25-32” based on improved FR4**



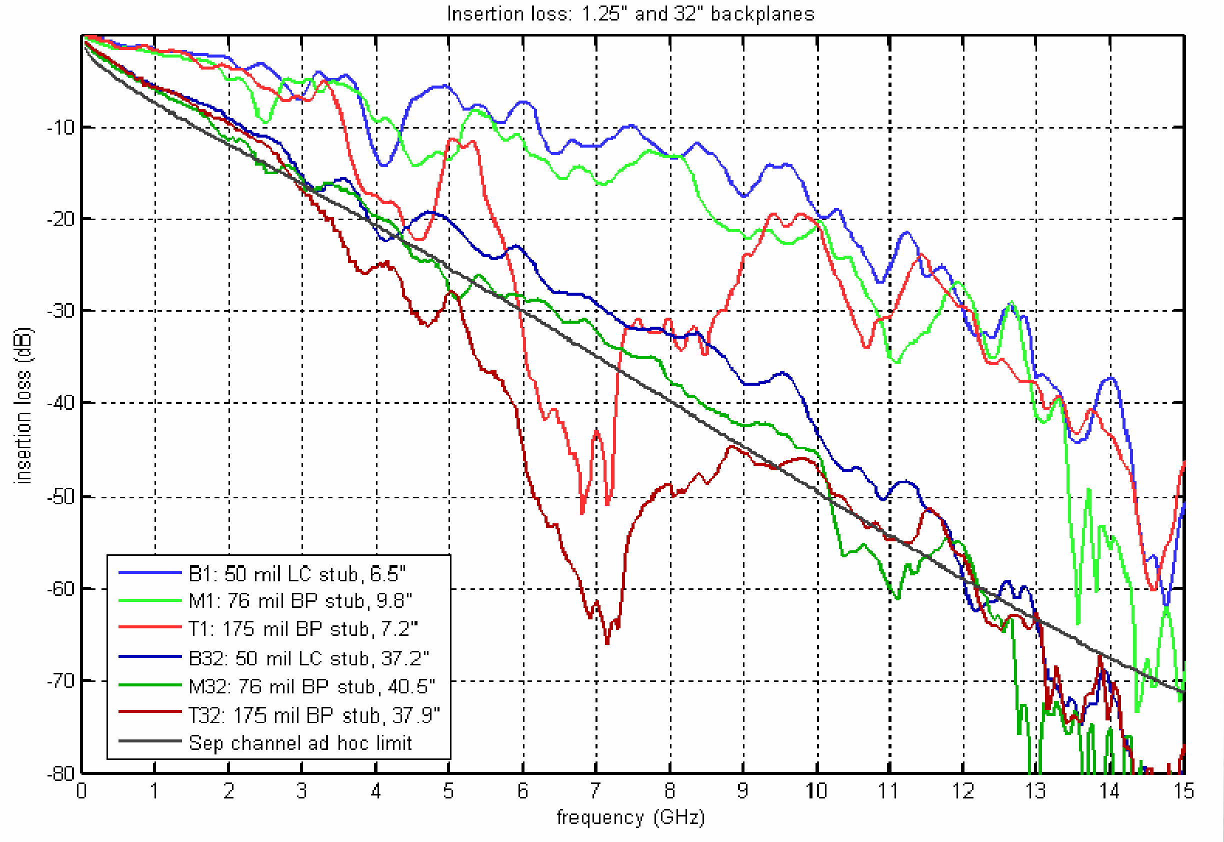
Ref. 6

BROADCOM. Connecting Everyone

BROADCOM. Connecting everything

IEEE Intel Channels

- **Insertion Loss for several ATCA Backplanes**
  - **Channel length 1.25-32” based on improved FR4**

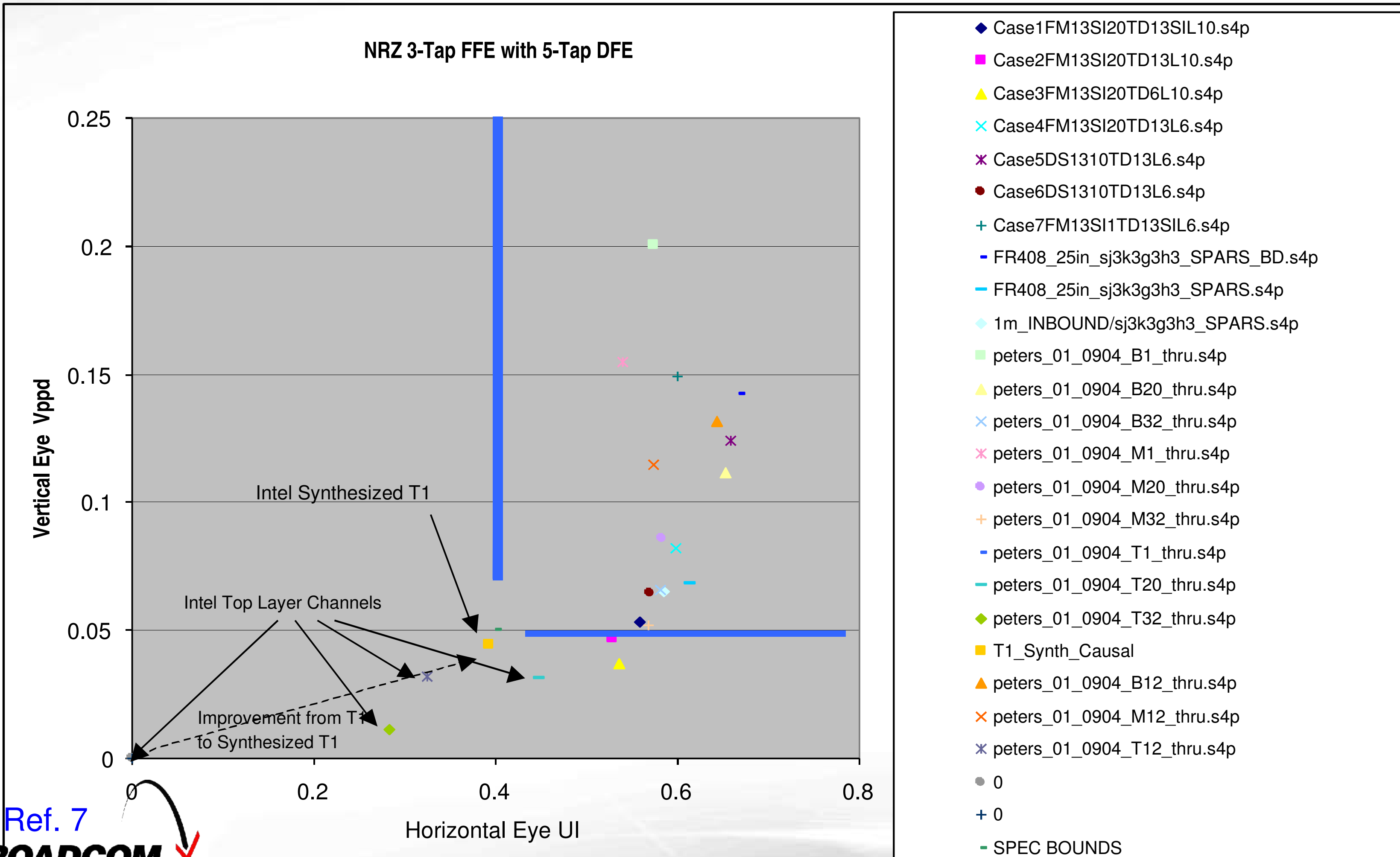


Ref. 6

BROADCOM. Connecting Everyone

BROADCOM. Connecting everything

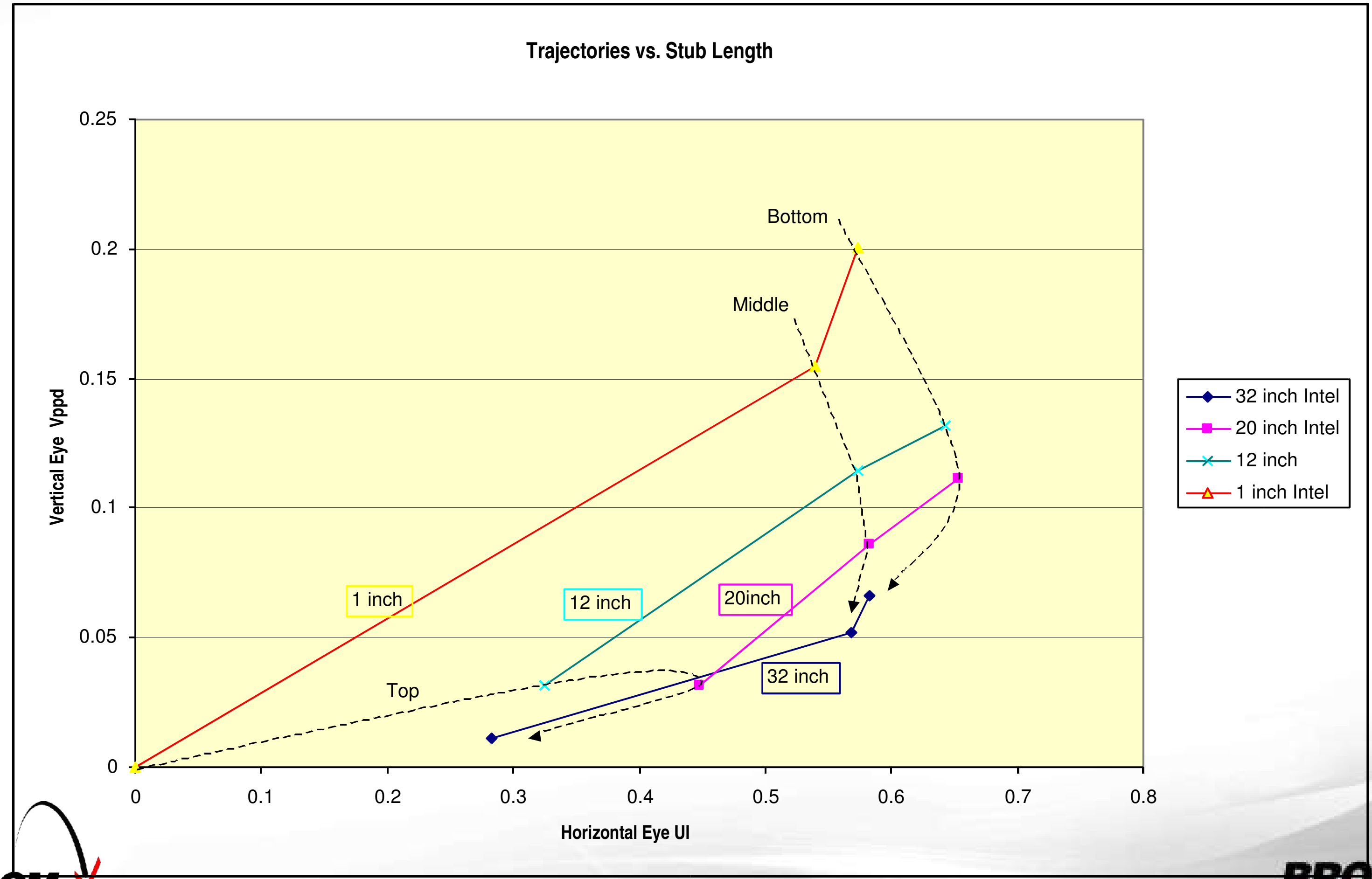
Vertical and Horizontal Eye Opening



Ref. 7

Eye Opening As Function of Channel and Stub Length

- **Stub length is the biggest obstacle for operating faster!**

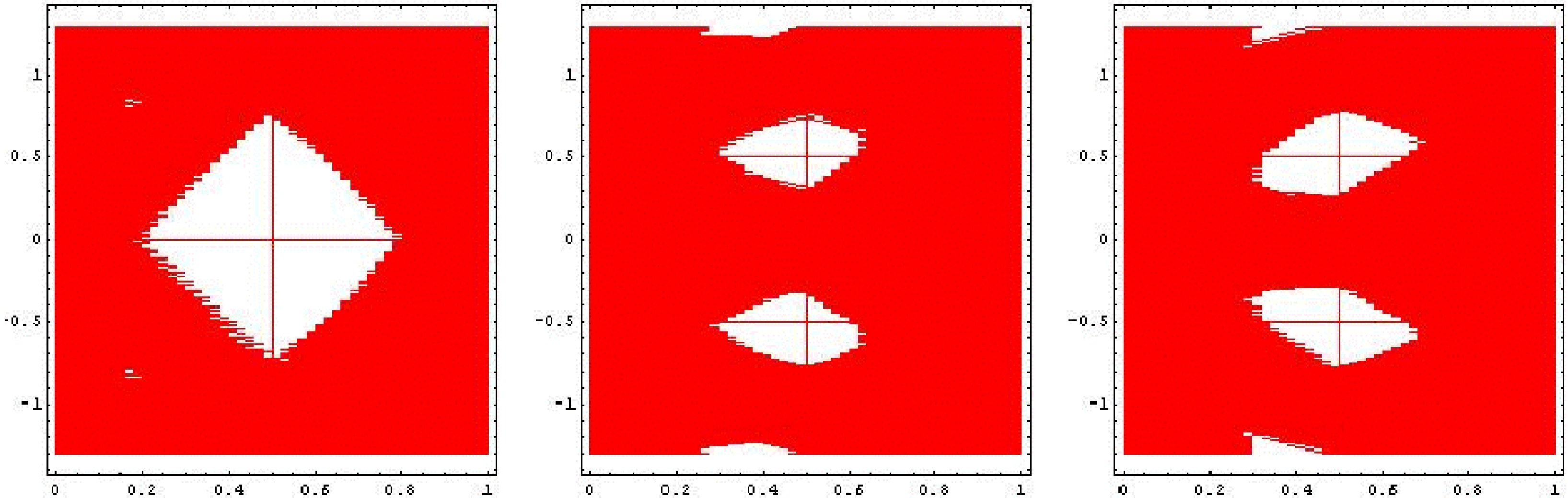


Ref. 7

Eye Margin For Tyco #1

- **Good channel, Eq with 4 Tap TX FIR and 5 Tap RX DFE.**

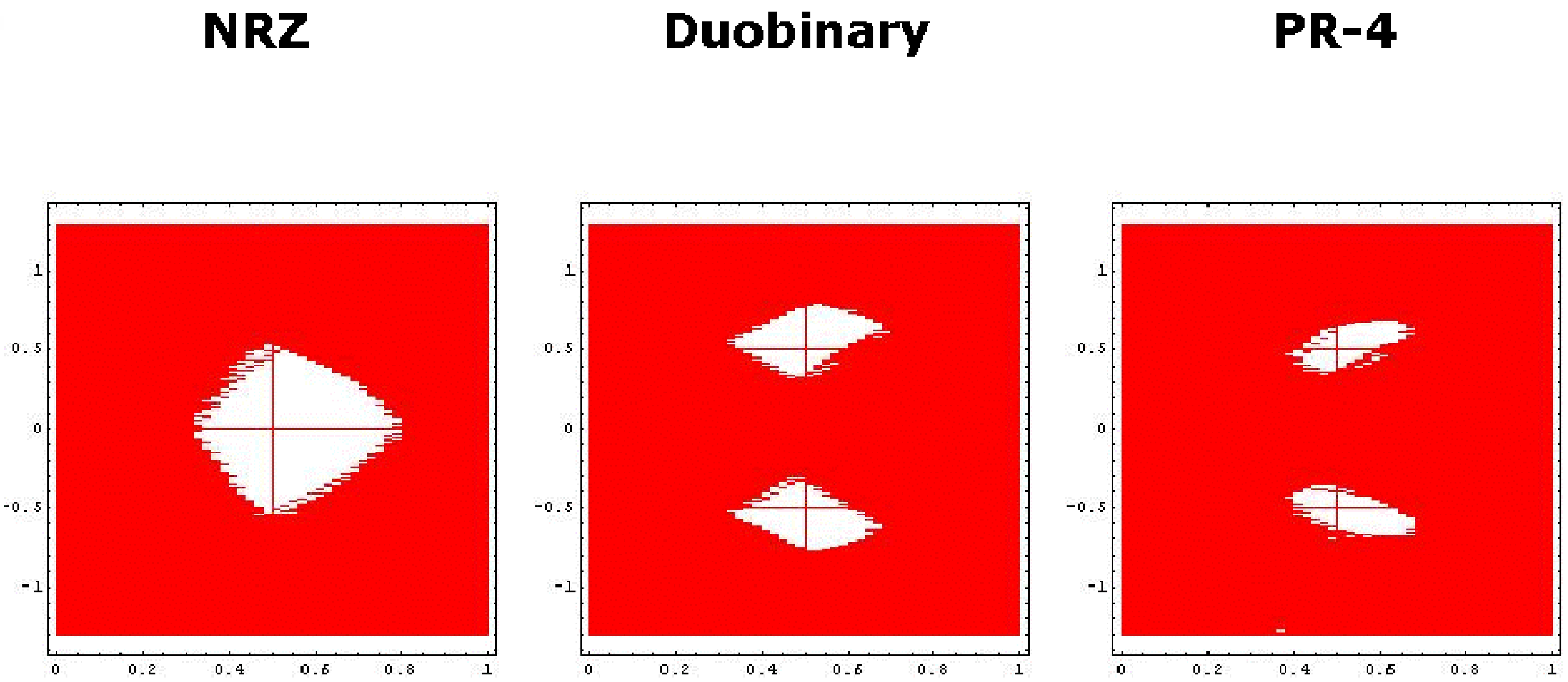
NRZ | **Duobinary** | **PR-4**



Ref. 4

Eye Margin For Intel B12

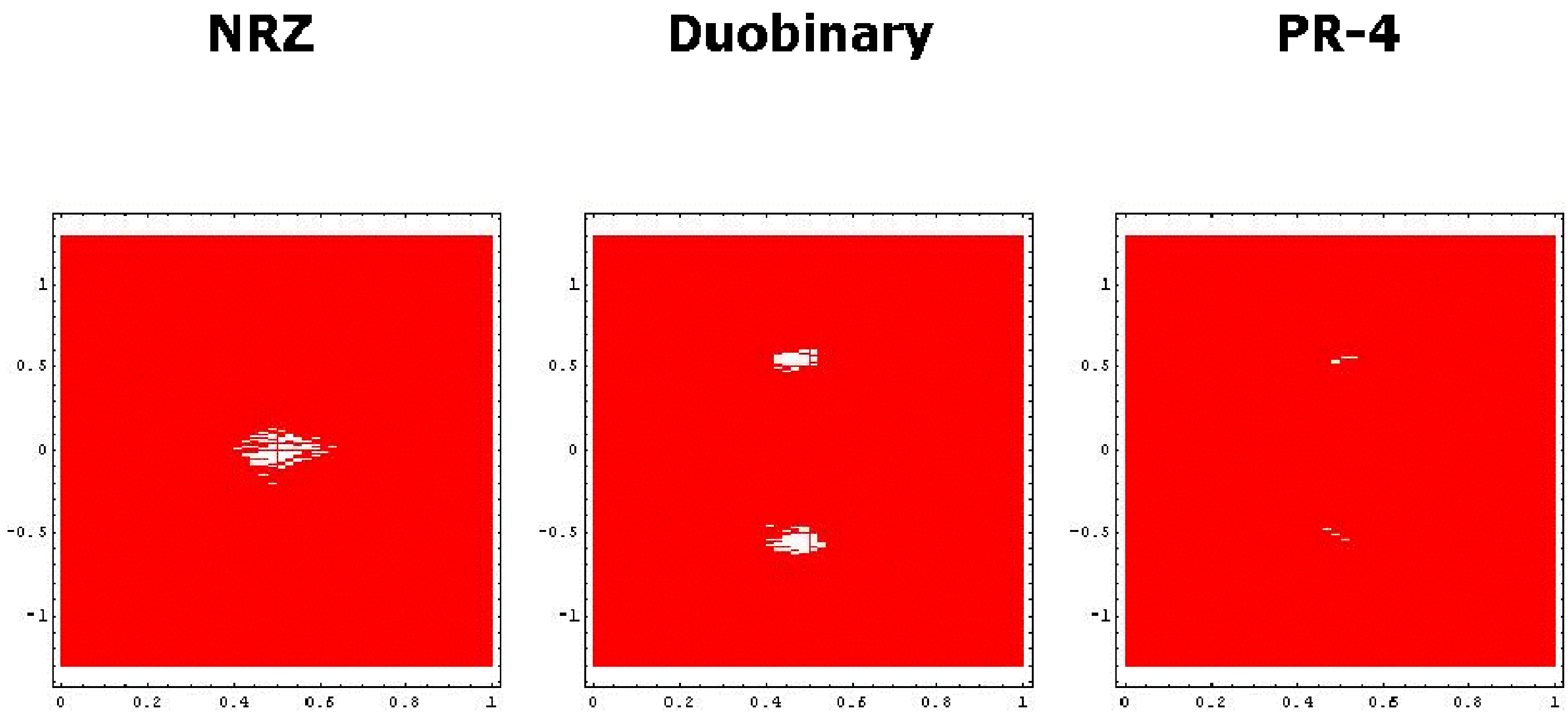
- **Good channel, Eq with 4 Tap TX FIR and 5 Tap RX DFE.**



Ref. 4

Eye Margin For Intel T1

- **Bad channel, Eq with 4 Tap TX FIR and 5 Tap RX DFE.**
  - **No satisfactory solution with three signaling scheme below.**



Ref. 4

BP Summary

- IEEE BP is scheduled to go to sponsor ballot by Jan 2006 and become about 6 month later.
- Methodology developed in IEEE Backplane Ethernet can assist development of SAS 6 Gig.
  - After more than 6 month of studying IEEE BP group selected NRZ signaling over PAM, PR2, and PR4.
- IEEE BP implementation based on NRZ signaling currently support
  - 1.25 Gbaud 1-lane (8B10B)
  - 3.125 Gbaud 4-lane (8B10B)
  - 10.3125 Gbuad 1-lane (64/66B)
- SAS Auto-negotiation can leverage 802.3ap DME to overcome OOB limitations.

BROADCOM. Connecting Everyone

BROADCOM. Connecting everything

DFE Error Propagation

- DFE are one of the common type of filter to equalize FR4 channels.
- DFE due to their feedback nature can propagate the error especially when the coefficient are large.
- CRC-32 can protect up to burst error of 3.
  - A 2 Tap DFE can generate burst of 3 error
  - A 4 Tap DFE can generate burst error >4.
- Three options exist for SAS 2.0
  - FEC probably too much latency
  - Add an additional CRC-8 to get total of 40 bits
  - Channel complexity is at level that DFE Tap <=2.

BROADCOM Connecting Everyone

BROADCOM Connecting everything

An Eye Toward SAS 2.0 Requirements

- Over-designing has an associated cost and power
  - Especially in SFF disk drives
- 1st step Defining the Channel
  - 26” vs 40”
  - Connector
  - FR4 type
  - Manufacturing practice and / or design role to manage stub length
  - Is there a need to operate an existing JBOD or Server at 6 Gig?
- OIF CEI 6 Gig LR may be overkill for 6 Gig SAS
- IEEE BP gives us path to 12 Gig SAS

References

1. IEEE Draft 802.3ap Rev. 0.9 Ethernet Operation Over the Backplane, Apr 20, 2005.
2. Robert Brink, 10GBASE-KR transmitter Compliance Methodology proposal, brink_01_0505.pdf.
3. Pat Thaler, Transceiver friendly auto-negotiation signaling for 802.3ap, thaler_01_1104.pdf.
4. Spagna and Altmann, IEEE 802.3ap Simulation Results for 10Gb Serial Links, altman_02_0305.pdf
5. Peters, Chen, Gong, Cai, and Austin, ATCA Channel data for Backplane Ethernet Task Force, peters_01_0904.pdf.
6. John D'Ambrosia, Proposed Changes to the SDD11/SDD22 Return Loss mask, dambrosia_02_0904.pdf.
7. Steve Anderson, Signaling Method Performance Results, anderson_01_0305-1.pdf.
8. IEEE backplane Ethernet Website http://ieee802.org/3/ap

Acknowledgment's

1. Adam Healey, Chair IEEE 802.3ap (Backplane Ethernet).
2. Steve Anderson from Xilinix.

BROADCOM Connecting Everyone

BROADCOM Connecting everything