100GbE Electrical Backplane / Cu Cabling Call-For-Interest

IEEE 802.3 Working Group

Dallas, TX

November 9, 2010

Objectives for this meeting

- To *measure the interest* in starting a study group for *100GbE Electrical Backplane and Cu Cabling*

- We don't need to
  - Fully explore the problem
  - Debate strengths and weaknesses of solutions
  - Choose any one solution
  - Create project documentation
    (i.e. PAR, 5 Criteria, Project Objectives)
  - Create a standard or specification

- Anyone in the room may speak / vote

- RESPECT… give it, get it

Agenda

- Presentations
  - “The Need”, Howard Frazier, Broadcom.
  - “Technical Viability”, Adam Healey, LSI.
  - “Why Now?”, John D’Ambrosia, Force10 Networks.
- Q&A
- Straw Polls
- Future Work

The Need

Presented by
Howard Frazier
Broadcom

IEEE 802.3 Working Group
Dallas, TX
November 9, 2010

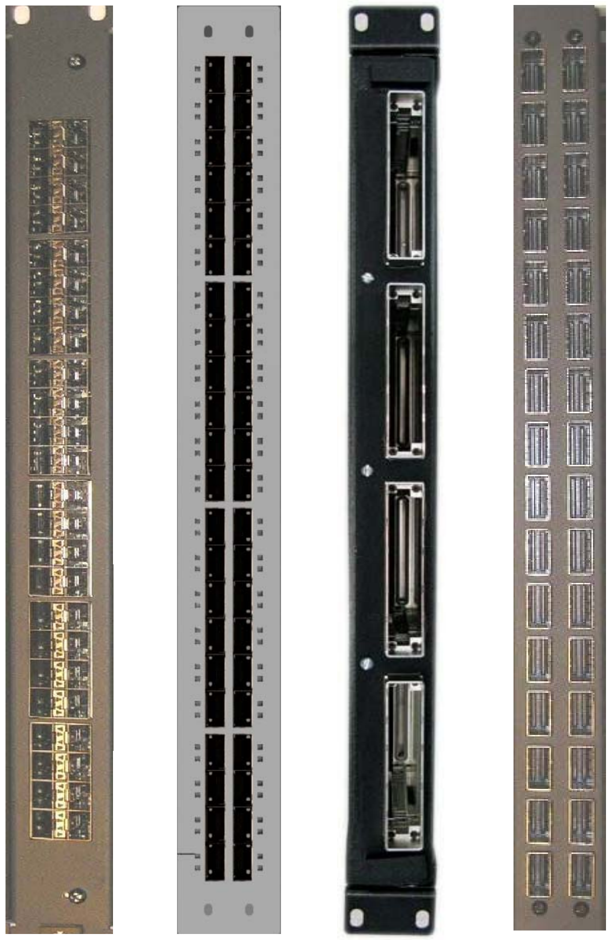
Historical perspective

- 10GBASE-KR: 10Gb/s serial backplane operation
  - Blade servers justified the effort
  - Defined a backplane channel model
- During IEEE P802.3ba project
  - 10GBASE-KR: basis for 40 / 100GbE PMDs
    - 40GBASE-KR4
    - 40GBASE-CR4
    - 100GBASE-CR10
  - No solution for 100GbE backplane

Speed and capacities

- Higher speed electrical signaling
  - Other industry bodies exploring 20 to 28Gb/s per lane
  - Ethernet has a historical preference for x4 interfaces
  - Note – using 4 x 25Gb/s as an example interface for this presentation
- Front panel I/Os
  - 10GbE, 40GbE, 100GbE Options
  - Higher density offerings at all speeds
    - Front panel can provide 100's of Gigabits of I/O

Front panel I/O driving backplane capacities



a b c d

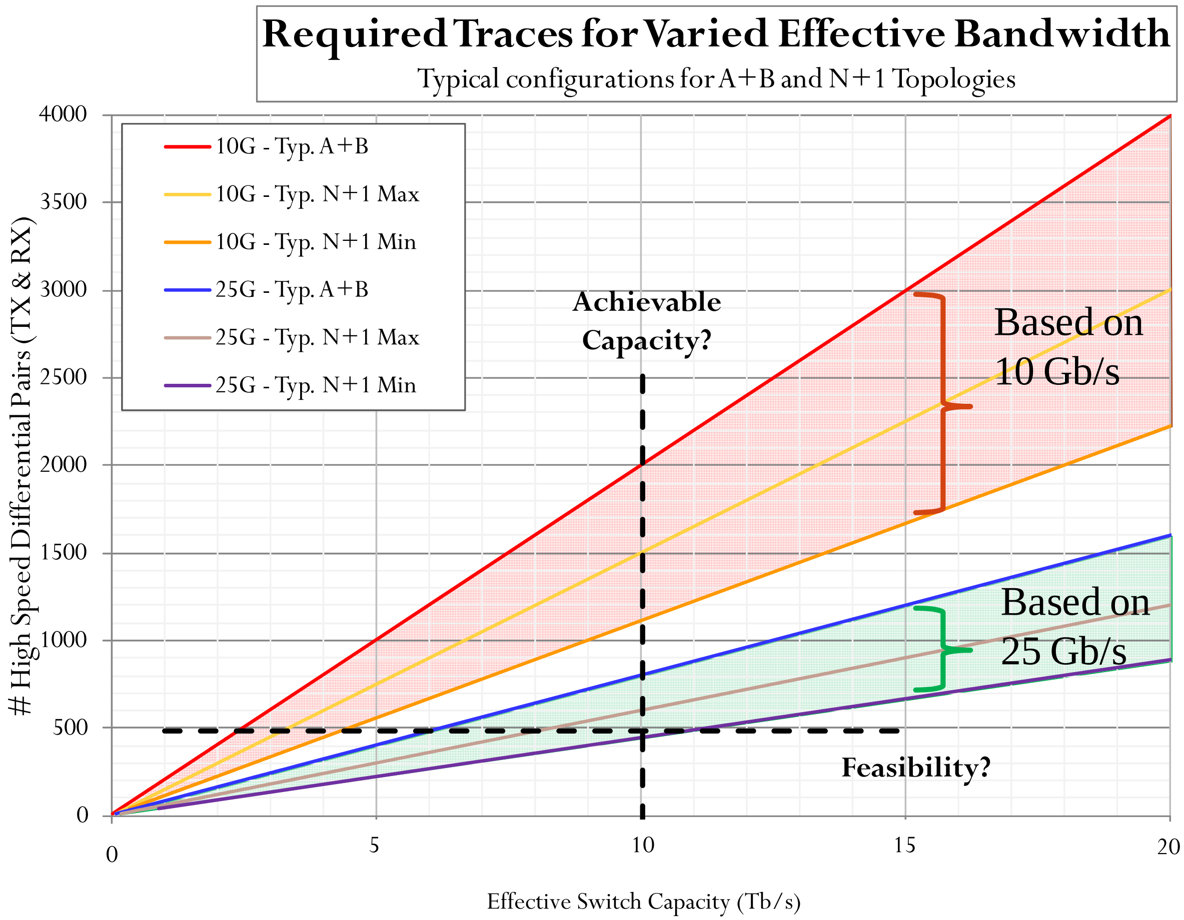
Line card illustrations

a. 48 ports SFP+ @ 10GbE = 480Gb/s
b. 44 ports QSFP @ 40GbE = 1.76 Tb/s
c. 4 ports CFP @ 100GbE= 400 Gb/s
d. 32 ports CXP@ 100GbE= 3.2 Tb/s

Potential backplane bandwidth capacities

- 8 Line Cards: 3.2 Tb/s to 25.6 Tb/s
- 14 Line Cards: 5.6 Tb/s to 44.8 Tb/s

Front panel I/O driving backplane capacities



Note – Dashed black lines represent examples for consideration

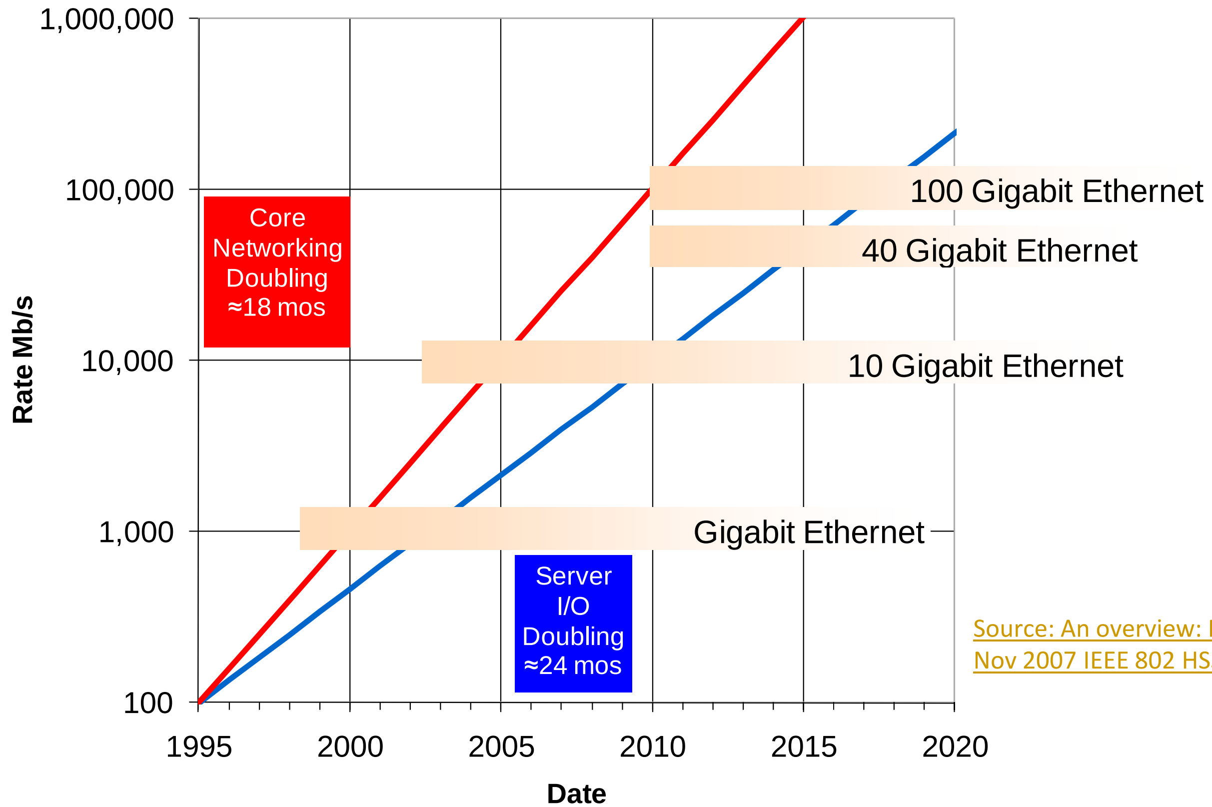
<u># of signals will impact</u>

- Feasibility
  - Support target backplane bandwidth?
  - Connector density (Pair / inch)?
  - Can it be built?
- Cost
  - # of connector pins
  - # of PCB layers

Server I/O BW drivers

- Higher system processing capability
  - Multi-core processors
  - Higher speed memory, systems bus, and next gen. process technologies
- Server virtualization
  - Consolidation of multiple logical servers in a single physical server
- Converged networking and storage
  - Multiple I/O connections converging to single connection with fabric virtualization
- Clustered servers
  - Scientific, financial, oil/gas exploration, engineering workloads
- Internet applications
  - IPTV, Web 2.0

Server bandwidth trends



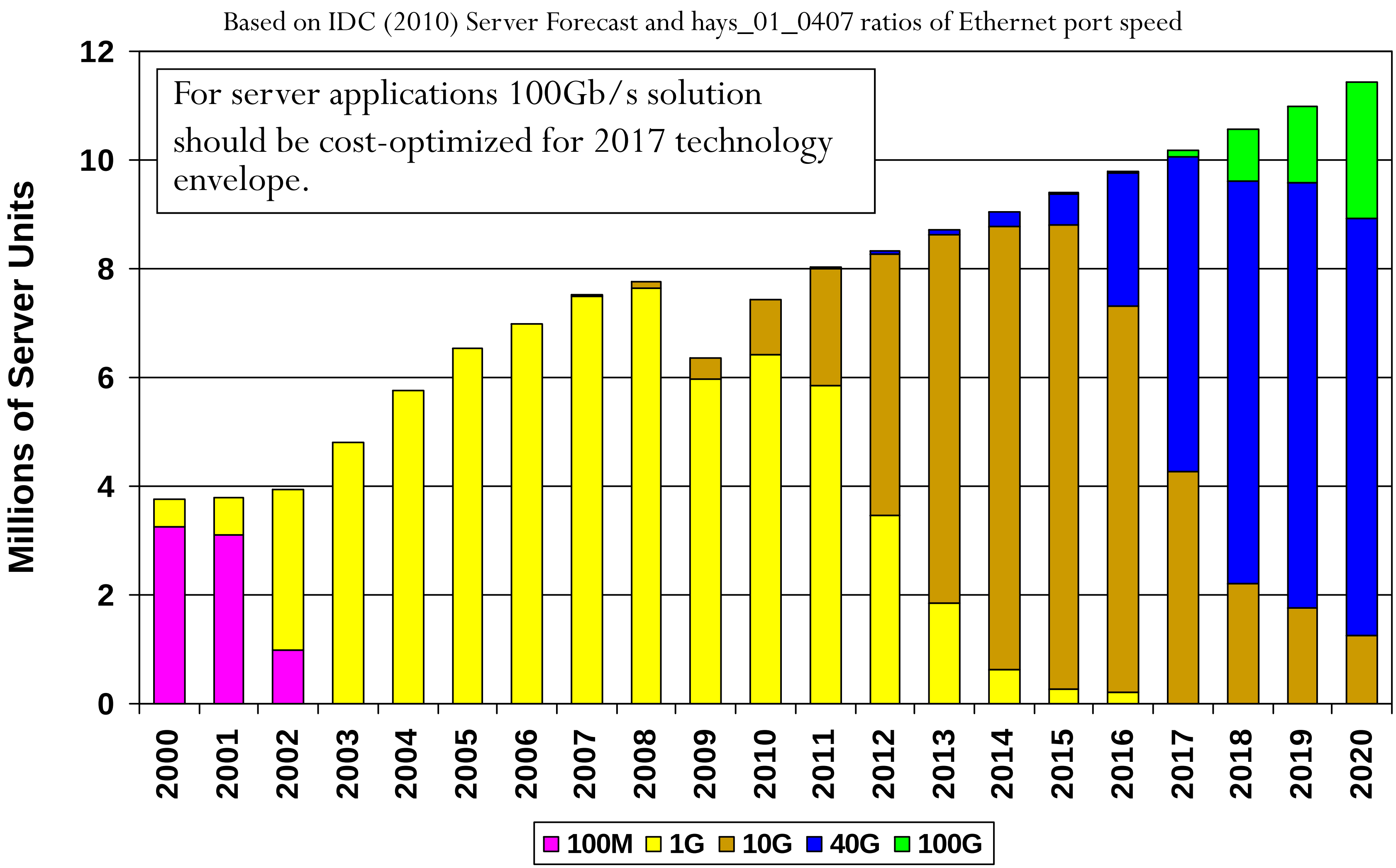
Source: An overview: Next Generation of Ethernet – Nov 2007 IEEE 802 HSSG tutorial

Server I/O bandwidth
- Doubles ≈ every 24 mos
- 100G Need: 2017

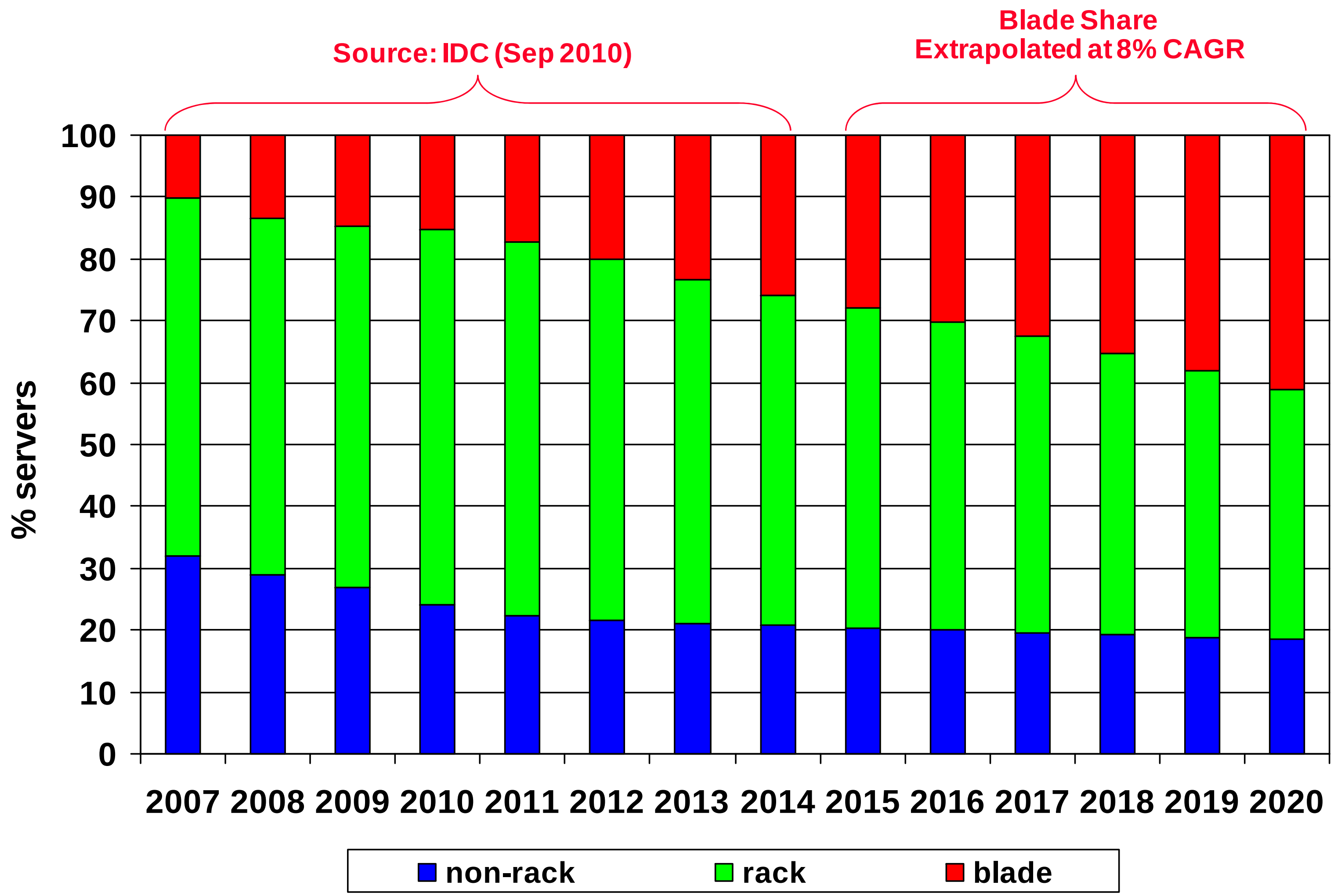
Development cycles
- Standards: 2-3 years
- System and silicon: 2-3 years
- Expect 5 years from start of 802.3 project to first products

Next Gen blade systems
- Forward migration path:
  - 10G → 40G → 100G
- Backplanes supporting 100G will be shipped prior to 100G Servers
- Standards for new speeds needed

x86 servers by Ethernet connection speed (2010 forecast)

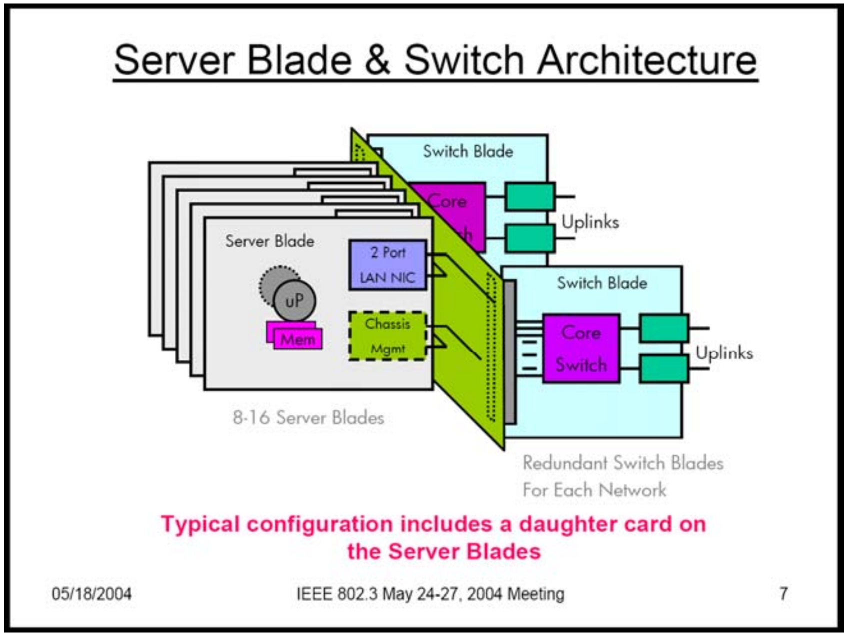


Trends in server form factors



Backplane Ethernet & blade server architectures

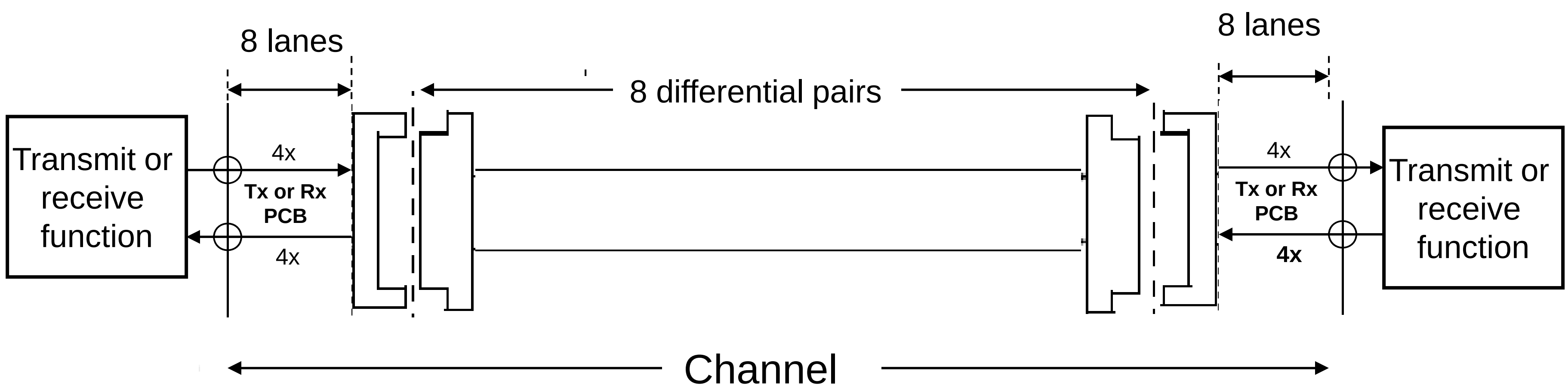
- IEEE Std 802.3™-2008 defines GbE and 10GbE operation over a modular platform backplane (1 m objective)
  - 1000BASE-KX (GbE)
  - 10GBASE-KX4 (10 GbE, 4 x 3.125 GBd)
  - 10GBASE-KR (serial 10 GbE)
- Blade servers: 2nd Gen backplanes
  - Based on 10GBASE-KX4 architecture...
  - ...but satisfy 10GBASE-KR channel requirements
  - IEEE Std 802.3ba™-2010 introduced 40 Gb/s operation on backplanes: 40GBASE-KR4 (4 x 10.3125 GBd)
- Blade servers: 3rd Gen backplanes
  - Backwards compatibility needed
  - 100GbE must support 4 lane approach



Source: Koenen, "Channel Model Requirements for Ethernet Backplanes in Blade Servers", May 2004.

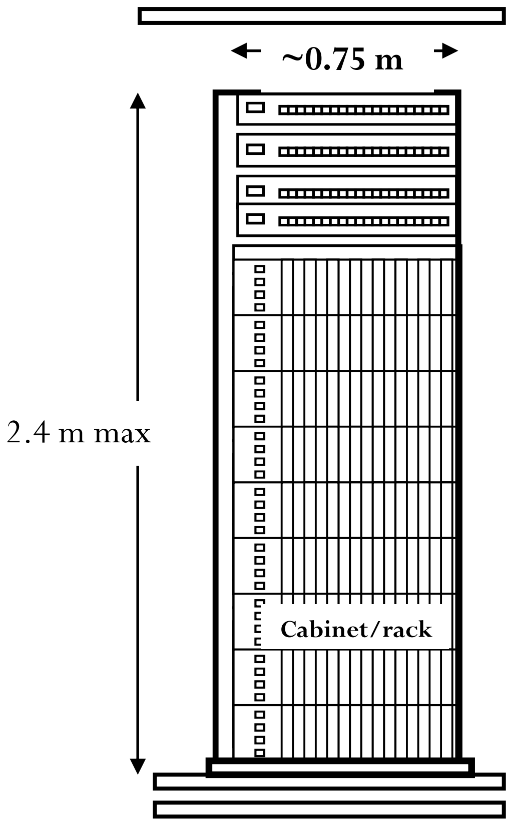
Cu Cabling: Why 4x25Gb/s vs 10x10Gb/s?

- 8 lanes/pairs versus 20 lanes/pairs
  - Smaller form factor enables higher port densities.
- Cost factors
  - Fewer cable assembly pairs reduces cost.
    +20-pair assembly ~1.5 times cost of 8-pair assembly.
  - Reduction in number of Tx/Rx lanes simplifies host routing.
  - Reduction in cable assembly pairs enables reduction in cable diameter.
  - Fewer pairs => higher levels of integration => higher port density
- Could define plug compatible port types to enable plug-and-play for 4x10Gb/s and 4x25Gb/s.



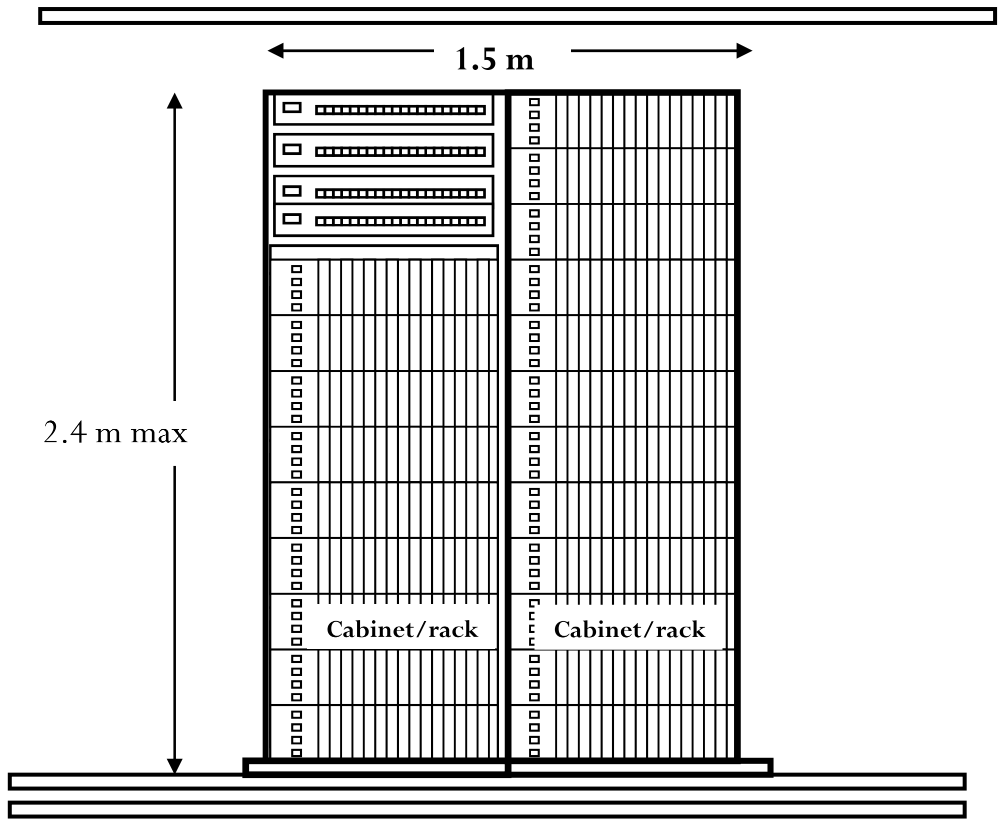
Media reach

Intra-rack configurations



- at least 3m addresses majority of configurations

Inter-rack configurations



- 3 to 7m addresses a meaningful portion of configurations

Summary

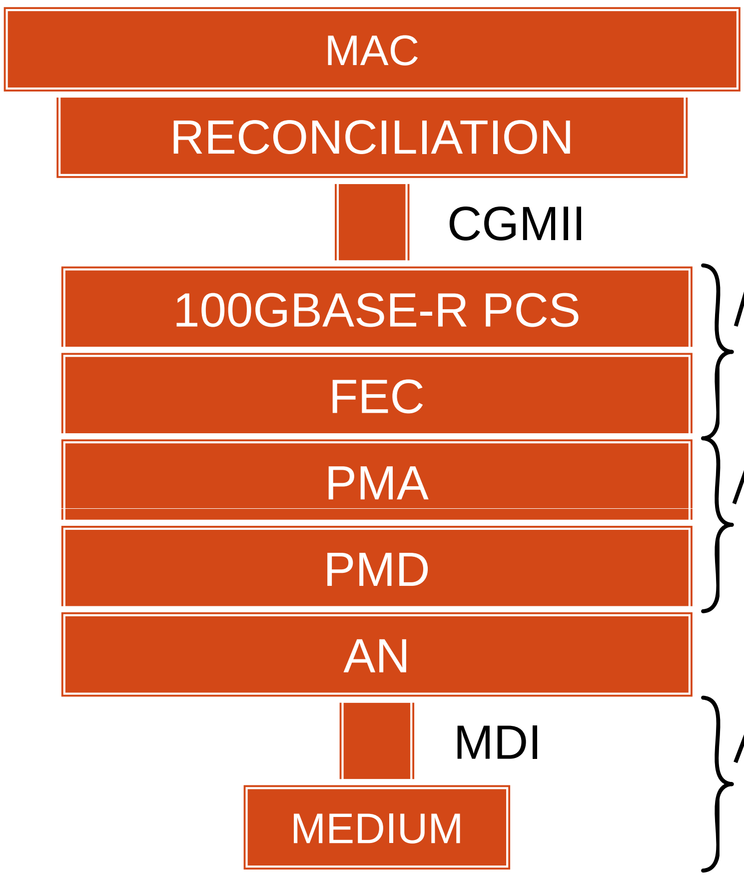
- Backplane application space
  - Front panel capacities are challenging backplane capacities
  - Server requirements
- Cu cabling
  - Intra-rack & inter-rack applications
  - Can leverage backplane solutions
  - Potential lower cost and wider application usage
- Timing Considerations
  - Network equipment backplanes and inter / intra-rack connections
  - Blade servers / backplanes & rack servers

Technical Viability

Presented by
Adam Healey
LSI Corporation

IEEE 802.3 Working Group
Dallas, TX
November 9, 2010

Potential enablers for more Gb/s/lane



Increase coding gain (trade-off against latency)

Equalization
Crosstalk cancellation
Modulation (e.g. M-PAM)

Lower loss printed circuit board traces
Lower loss cabling
Low noise connectors
Improved impedance control

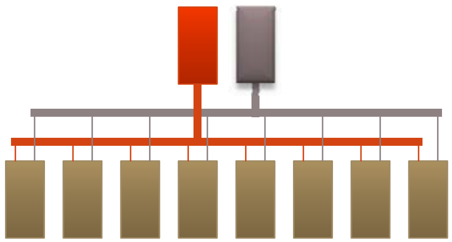
- Combine (a subset of) technologies to increase the per-lane throughput

Number of lanes and signaling rates

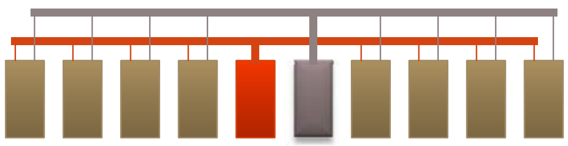
Number of lanes	Uncoded rate, Gb/s	Notes
10	10	State of the art, e.g. 100GBASE-CR10
5	20	Lower signaling rate per lane
4	25	Better alignment with existing physical architectures
2	50	Technology not yet available
1	100	Technology not yet available

- Spectrally efficient modulation reduces signaling rate per lane
- Opportunities to re-use the 100GBASE-R PCS
- 4 x 25 Gb/s will be used as the basis for subsequent examples

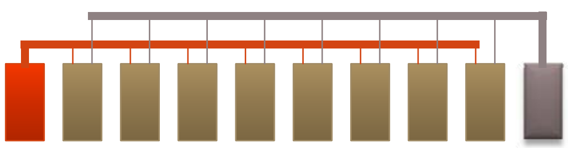
Survey of modular platform topologies



(a) Dual-star mid-plane



(b) Dual-star backplane, switches centered



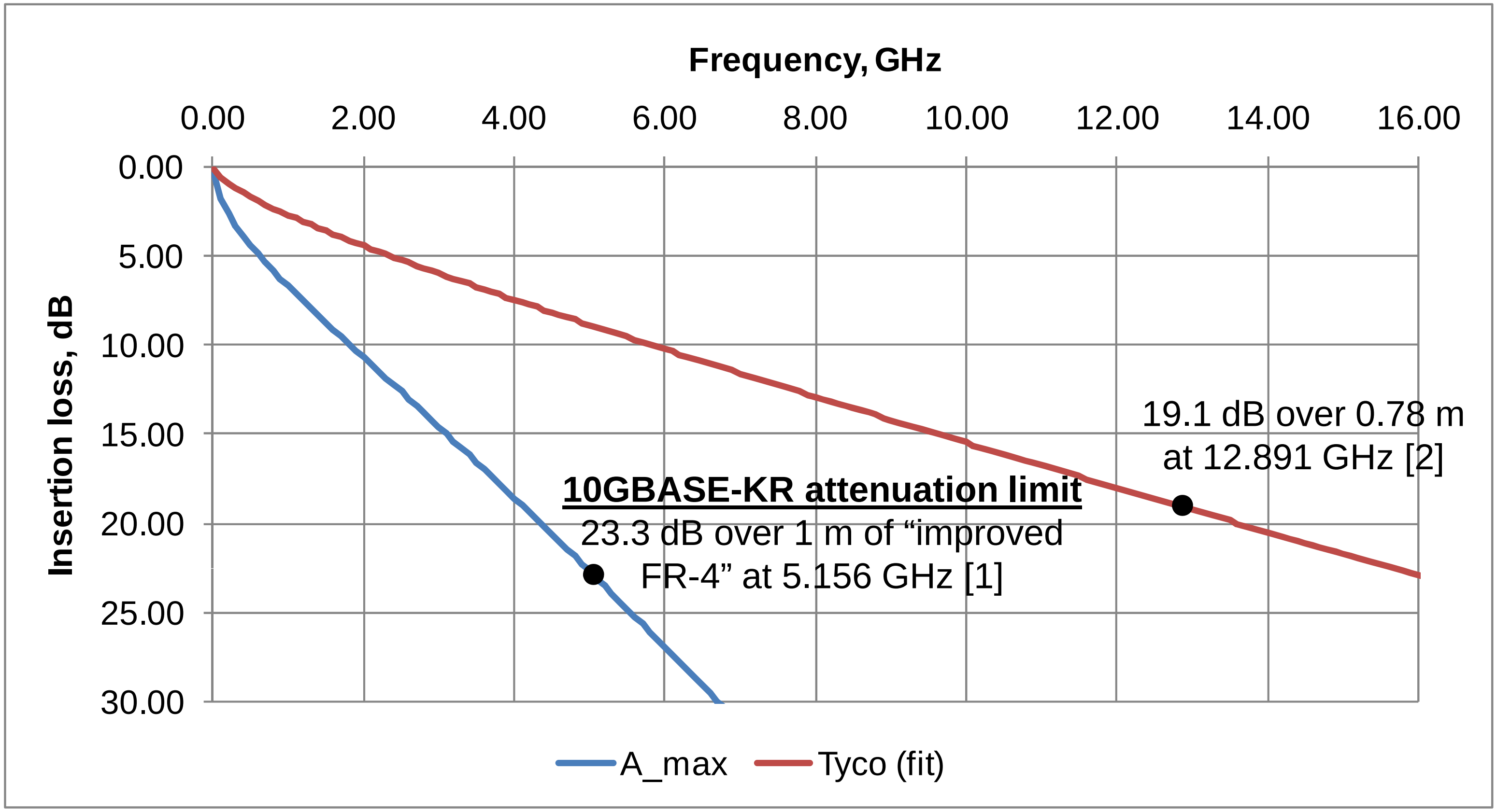
(c) Dual-star backplane, switches at edges

Source: Richard Mellitz, Intel

Topology	Distance, in.
Dual-star mid-plane	11 to 22
Torus-like	11 to 22
Dual-star backplane (switches centered)	12 to 26
Dual-star backplane (switches at edges)	22 to 39
Full-mesh	22 to 39

- Dual-star predominant in the server marketplace
- 10GBASE-KR targeted 1 m (~39 in.)

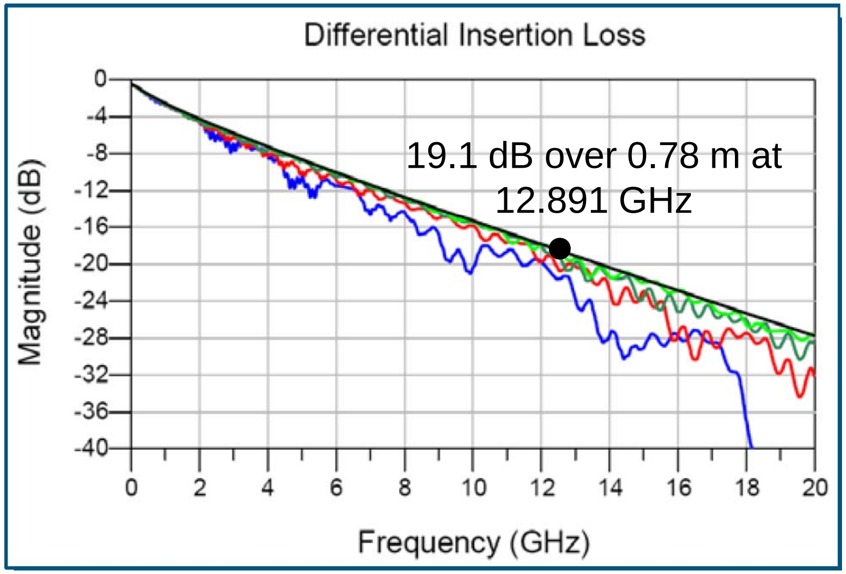
Backplane channel loss

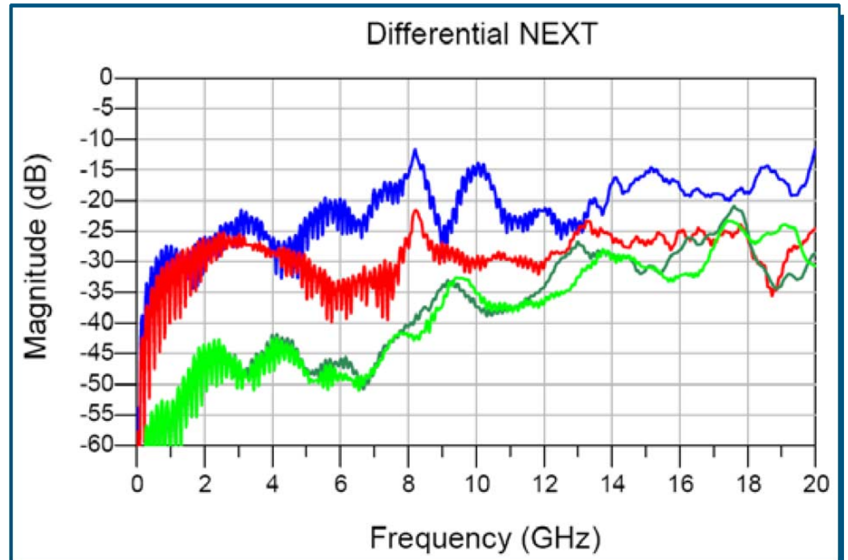


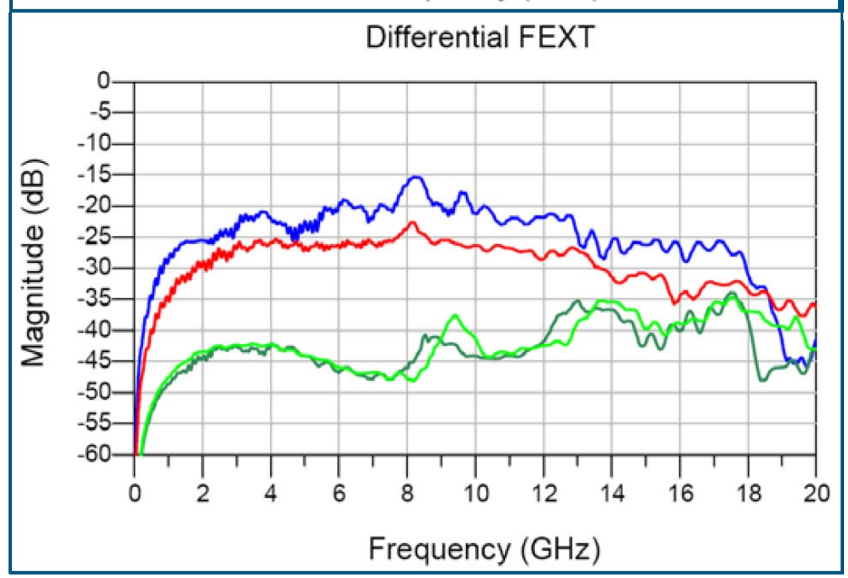
[1] IEEE Std 802.3TM-2008, 69B.4.2

[2] Source: Tyco Electronics, 0.51 m backplane, two 0.10 m line cards, two 0.035 m ideal connectors, 6-8-6 mil differential stripline, Nelco 4000-13SI dielectric, 1 oz copper

Backplane connector evolution





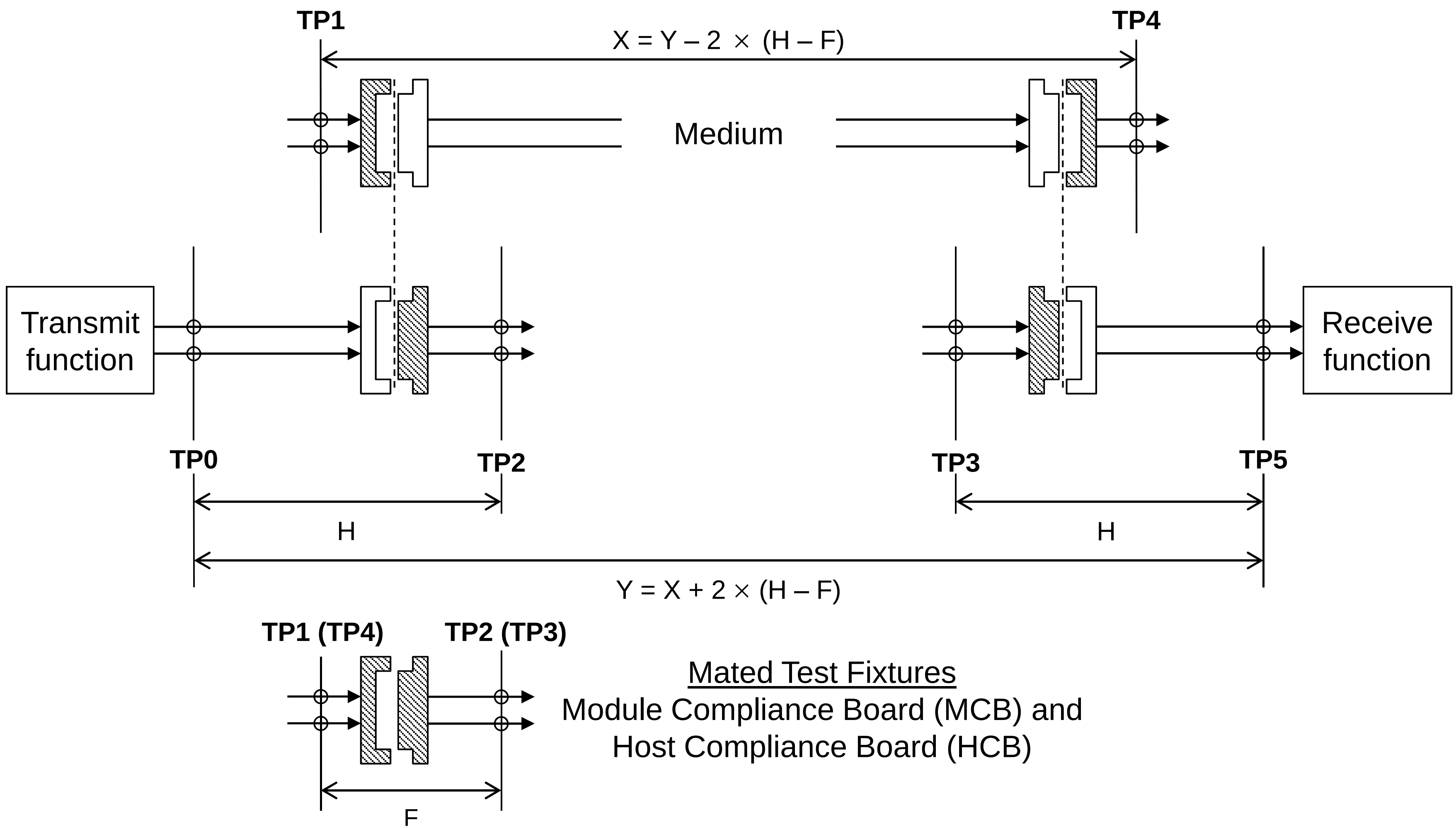


Gen1, 100 Ω connector (and footprints)
Gen2, 100 Ω connector (and footprints)
Gen3, 100 Ω connector (and footprints)
Gen3, 85 Ω connector (and footprints)

- Improvements in impedance and noise control

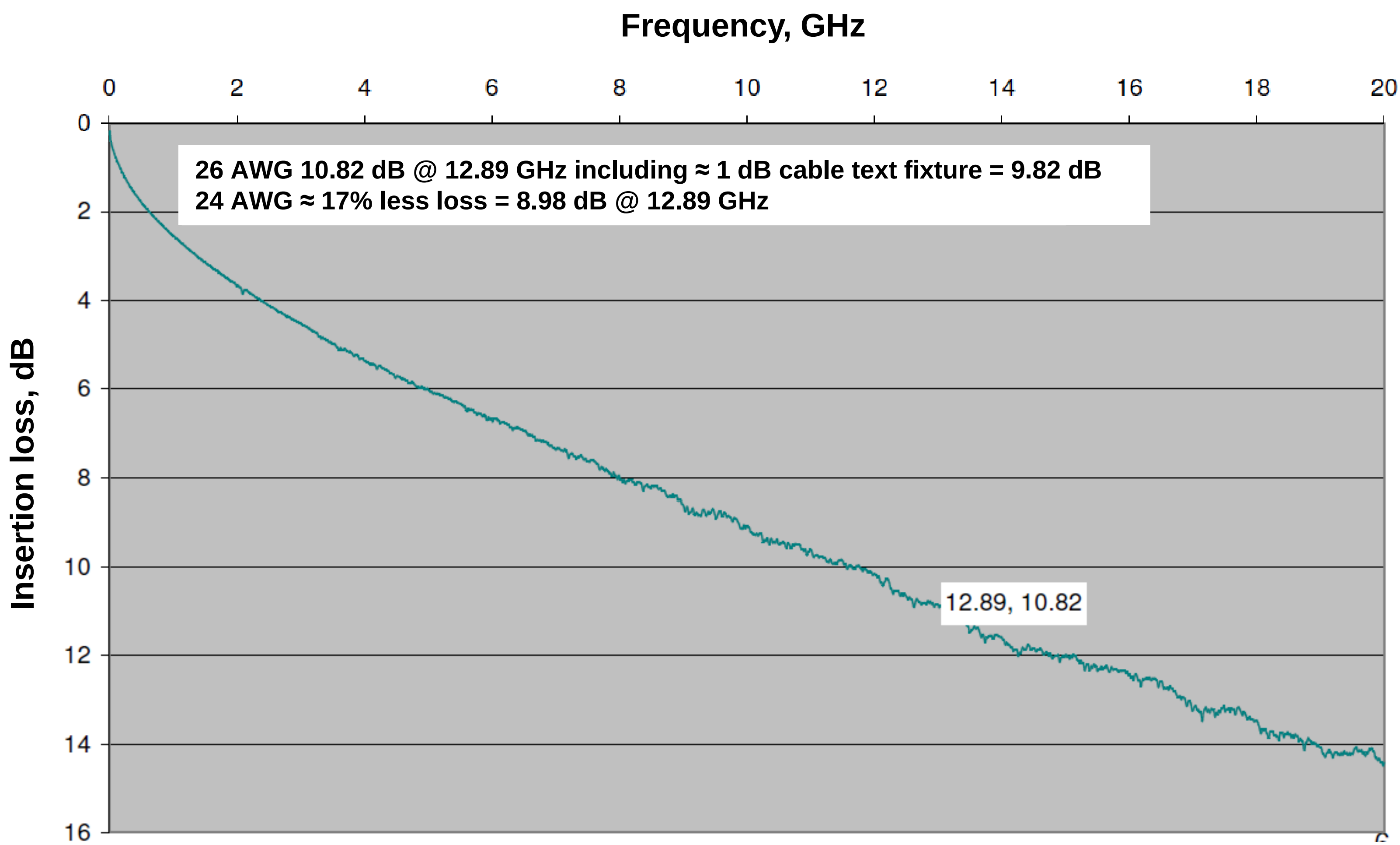
Source: Tyco Electronics

Passive cable assembly topology



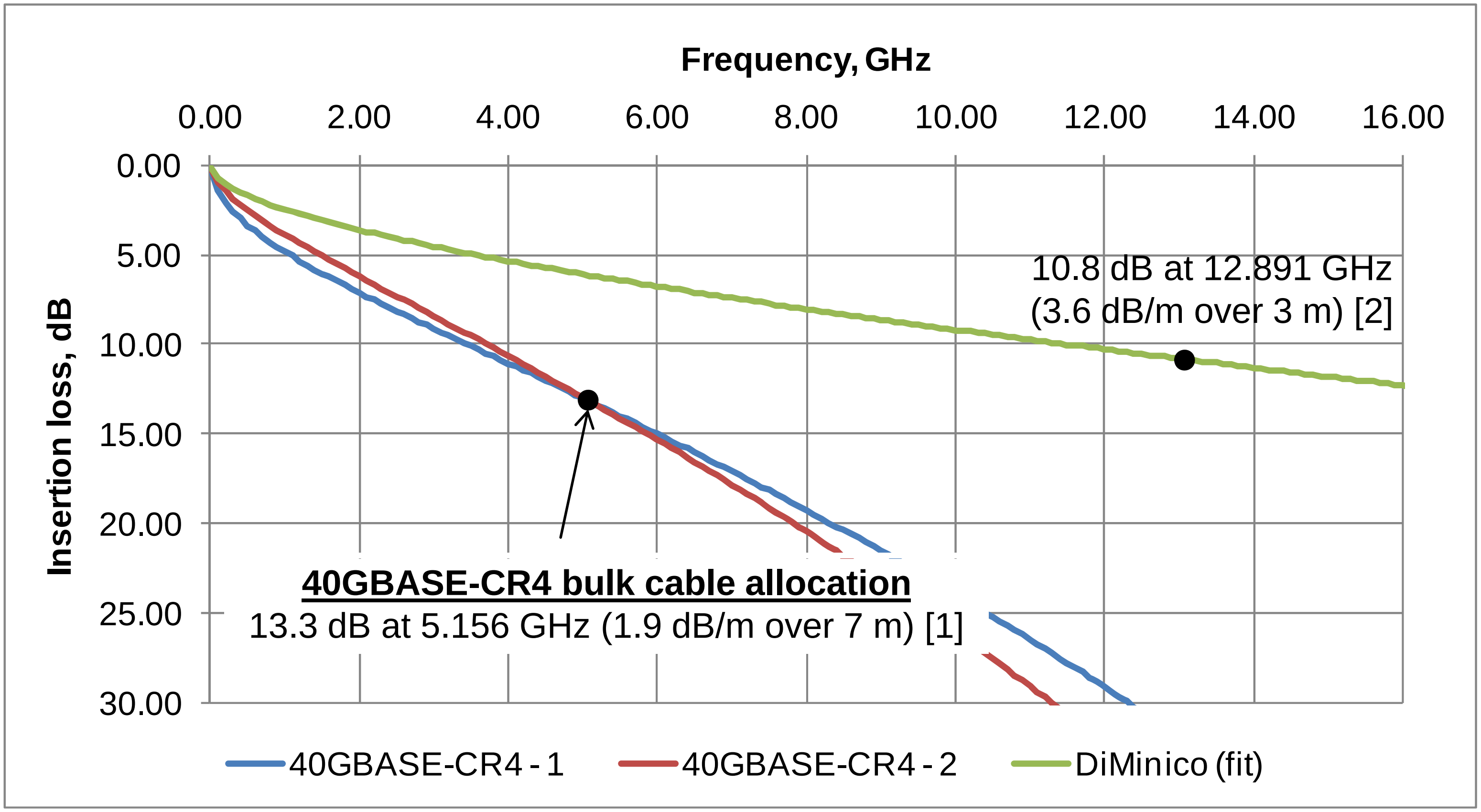
Mated Test Fixtures
Module Compliance Board (MCB) and
Host Compliance Board (HCB)

Measured bulk cable loss, 3 m



Source: Chris DiMinico, Leoni

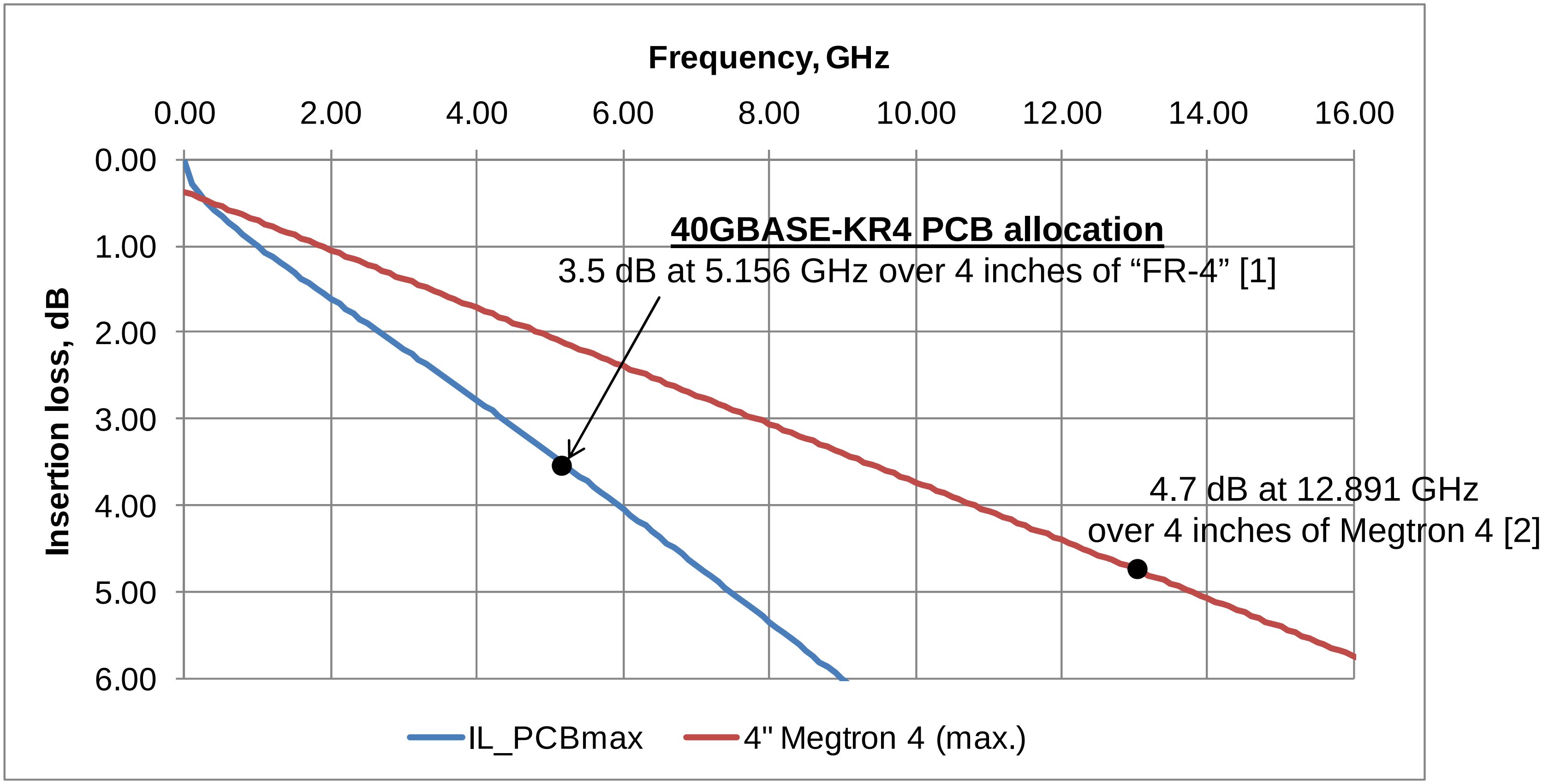
Twinaxial bulk cable insertion loss



[1] Derived from IEEE Std 802.3ba™-2010 Equations 85-19 (IL_{fitted}) and 85-34 (IL_{catf}) assuming that the insertion loss of connector, paddle card, and wire termination is 1.2 × sqrt(*f*/5.1563) where the units of *f* are GHz.

[2] Chris DiMinico (Leoni), 26AWG, includes test fixture loss (approximately 1 dB at 12.89 GHz)

Host PCB insertion loss



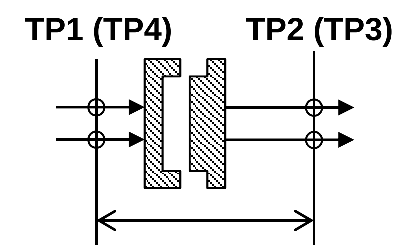
[1] IEEE Std 802.3ba^{TM}-2010, 85A.4.
[2] Marco Mazzini and Gary Nicholl (Cisco), "CEI-28G-VSR -Channel Distance and Loss Budget Requirements" Formulae for maximum insertion loss per inch extracted from channel measurements, W = 4 mils Megtron 4: $IL = 0.0838 f + 0.0944$.

Front-panel connector evolution

Parameter	40GBASE-CR4	Supplier A [2]	Supplier B [3]
Signaling rate, GBd	10.3125	25.0	25.0
Mated test fixture loss [1], dB	2.80	~6.4	~7.0
MCB PCB loss, dB	0.67	1.8	2.1
HCB PCB loss, dB	1.26	3.5	4.0
Extracted connector loss, dB	0.87	~1.1	~0.9
MDFEXT[4] ICN[5], mV	3.5	1.3	1.5
MDNEXT[4] ICN, mV	1.0	0.5	0.5

[1] Losses are defined at the fundamental frequency for the cited signaling rate.
[2] Source: Amphenol (http://www.ieee802.org/3/ad_hoc/OIF_VSR_liaison/public/AMPHENOL%20QSFP%20data.pdf)
[3] Source: Molex/Tyco Electronics (http://www.ieee802.org/3/ad_hoc/OIF_VSR_liaison/public/fromm_01_0610.pdf)
[4] MDFEXT = Multi-Disturber Far-End Crosstalk, MDNEXT = Multi-Disturber Near-End Crosstalk
[5] ICN = Integrated Crosstalk Noise, refer to IEEE Std 802.3ba™-2010 85.10.7.

- Data in this table represents solutions that are compatible with the QSFP mating interface



Loss budget examples

	40GBASE-CR4 signal integrity		"Next generation" signal integrity	
Uncoded rate, Gb/s	10.0	25.0	25.0	25.0
Line code	NRZ	4-PAM	4-PAM	NRZ
Signaling rate, GBd	10.3125	12.8913	12.8913	25.7813
SNR for BER ≤ 10^{-12}, dB [1]	17.0	26.6	26.6	17.0
Cable length, m	7	7	7	3
Host TX PCB (4") [2], dB	3.50	4.33	2.54	4.70
TX Connector, dB	2.07	2.31 [3]	1.41 [4]	2.00
Bulk cable, dB	13.30	16.42	13.68 [5]	10.82
RX Connector, dB	2.07	2.31 [3]	1.41 [4]	2.00
Host RX PCB (4"), dB	3.50	4.33	2.54	4.70
Total insertion loss, dB	**24.44**	**29.70**	**21.58**	**24.22**

[1] Assumes fixed transmitter peak-to-peak differential output voltage.
[2] Losses are defined at the fundamental frequency for the cited signaling rate.
[3] Derived as 2.07 × sqrt(6.4453/5.1563)
[4] Derived as 2.00 × sqrt(6.4453/12.8913)
[5] Derived as (7/3) × 7.06 × 0.83 where 0.83 is the reduction in loss for 24AWG cabling relative to 26AWG

Other industry efforts

Organization	Project	Notes
Optical Internetworking Forum (OIF) [1]	CEI-28G-(V)SR	• 19.9 to 28.05 GBaud/lane • Chip-to-chip, chip-to-module
	CEI-25G-LR	• 19.9 to 25.8 GBaud/lane • Backplane
Infiniband Trade Association (IBTA)	EDR	• 25.78125 GBaud/lane • Passive copper, active cables
INCITS T11 Fibre Channel	32GFC	• 28.05 GBaud (single lane) • Chip-to-module, multimode and single-mode fiber

[1] Refer to the proceedings of the IEEE 802.3 Ethernet Working Group OIF CEI-28G-VSR liaison response ad hoc (http://www.ieee802.org/3/ad_hoc/OIF_VSR_liaison/index.html).

- Broader industry movement towards higher throughput per lane

Summary

- Rich selection of technologies for a Study Group to explore
- Representative of a broader industry movement

Why Now?

Presented by
John D'Ambrosia, Force10 Networks

IEEE 802.3 Working Group
Dallas, Tx
November 9, 2010

Why Now?

- There is no 100GbE backplane solution, but…..
  - Networking equipment - faceplate capacities are challenging current backplane solutions
    - o 10G I/O (SFP+, 10GBASE-T) densities are growing
    - o Introduction of 40GbE / 100GbE & related MSAs
  - Blade Servers
    - o 40GbE Backplanes need to support 100GbE
      - 40GbE backwards compatibility needs
      - 100GbE backplane design requirements
- 100GbE Cu Interfaces narrower than 100GBASE-CR10
  - May leverage a 100GbE backplane solution
  - Could enable multiple benefits – backwards compatibility with 40GbE, higher port density, and lower cost

Supporters

- Rick Rabinovich, Alcatel-Lucent
- Steve Trowbridge, Alcatel-Lucent
- Mike Li, Altera
- Shawn Searles, AMD
- Brian Kirk, Amphenol TCS
- Greg McSorley , Amphenol
- Brad Booth, Applied Micro
- Matt Brown, Applied Micro
- Rita Horner, Avago
- Brian Misek, Avago
- Charles Moore, Avago
- Jeff Slavick, Avago
- Yakov Belopolsky, Bel Stewart
- Paul Kish, Belden
- Wael Diab, Broadcom
- Howard Frazier, Broadcom
- Ali Ghiasi, Broadcom
- Arthur Marris, Cadence
- Hecham Elkhatib, Cinch
- Beth Donnay, Cisco
- Joel Goergen, Cisco
- Mark Nowell, Cisco
- Bhavesh A. Patel, Dell
- Ghani Abbas, Ericsson
- Vittal Balasubramanian, FCI
- Chris Cole, Finisar
- John D'Ambrosia, Force10 Networks
- Kapil Shrikhande, Force10 Networks
- Uri Cummings, Fulcrum Microsystems
- Ryan Latchman, Gennum
- Francois Tremblay, Gennum
- Dave Koenen, HP
- Steve Carlson, High Speed Design
- Orlindo Savi, Hitachi
- Hidehiro Toyoda, Hitachi
- Jeff Lynch, IBM
- Pravin Patel, IBM
- Andy Moorwood, Infinera
- Andre Szczepanek, Inphi
- Dave Chalupsky, Intel
- Ilango Ganga, Intel
- Rich Mellitz, Intel
- Thananya Baldwin, IXIA
- Jerry Pepper, IXIA
- Jeff Maki, Juniper
- David Ofelt, Juniper
- Alan Flatman, LAN Technologies, UK
- Mike Bennett, LBNL
- Adam Healey, LSI
- Cathy Liu, LSI
- Marc Dupuis, Madison Cable
- Chris DiMinico, MC Communications/LEONI
- Oren Sela, Mellanox
- Howard Baumer, Mobius
- Galen Fromm, Molex
- Jay Neer, Molex
- Paul Vanderlaan, Nexans
- Fanny Minarsky, Octoscope
- Shimon Muller, Oracle
- Ron Nordin, Panduit
- Pat Diamond, Semtech
- George Eaton, Semtech
- Ed Cady, Siemon Company
- Valerie Maguire, Siemon Company
- George Zimmerman, Solarflare
- Sanjay Kasturia, Teranetics
- Karl Muth, Texas Instruments
- Iain Robertson, Texas Instruments
- Mike Fogg, Tyco Electronics
- Nathan Tracy, Tyco Electronics
- Jeff Lapak, UNH-IOL
- Frank Chang, Vitesse
- Ziad Hatab, Vitesse
- George Noh, Vitesse
- Atul Sharma, Volex
- Takeshi Nishimura, Yamaichi
- Marek Hajduczenia, ZTE

Q&A

Straw Polls

Call-For-Interest

- Should a Study Group be formed for "100GbE Electrical Backplanes and Cu Cable Interface"?

Y: 95 N: 0 A: 8

Participation

- I would participate in the "100GbE Electrical Backplane and Cu Cable Interface" Study Group in IEEE 802.3.

  Tally: 64

- My company would support participation in the "100GbE Electrical Backplane and Cu Cable Interface" Study Group in IEEE 802.3

  Tally: 43

Future Work

- Ask 802.3 to form 100GbE Electrical Backplane and Cu Cable SG on Thursday

- If approved
  - 802 EC informed of 100GbE Electrical Backplane and Cu Cable Interface SG on Friday
  - Email reflector / web page to be created
  - First 100GbE BP / Cu SG meeting, week of January 12 2011 IEEE 802.3 Interim.

Thank You!

Contributors

- Howard Frazier, Broadcom
- Elizabeth Donnay, Cisco
- Dave Chalupsky, Intel
- John D'Ambrosia, Force10 Networks
- Alan Flatman, LAN Technologies, UK
- Jeff Lynch, IBM
- Pravin Patel, IBM
- Ilango Ganga, Intel
- Rich Mellitz, Intel
- Adam Healey, LSI
- Marc Dupuis, Madison Cable
- Chris DiMinico, MC Communications/LEONI
- Jay Neer, Molex
- Mike Fogg, Tyco Electronics
- Nathan Tracy, Tyco Electronics