# Short Haul 10Gbps Ethernet Copper PHY Call for Interest

IEEE 802.3

Vancouver, Canada

November 14-18, 2005

http://www.ieee802.org/3/cfi/1105_1/10G_Short_Haul_CFI.pdf

# Outline and Presenters

- Introduction
  - Nick van Bavel, Vitesse
- Market Requirements and Potential
  - Dan Dove, DNS for HP ProCurve
- Cabling
  - Alan Flatman, LAN Technologies
- Technical Feasibility
  - Mike McConnell, KeyEye Communications
- Conclusion

# Additional Contributors

- Allan Nielsen, Tyco Electronics
- Chuck Hudson, HP Server Networking
- Ken Hodge, Brand-Rex
- Olindo Savi, Siemon

# Introduction and Overview

# Introduction

- Why Are We Here?
  - Only To Answer The Following Question

    "Should We Form A Study Group For A Short Haul 10Gbps Copper PHY?"

  - We Are Not Here To:
    - Try To Write The Objectives Of The Study Group
    - Try To Write The PAR Or The 5 Criteria
    - Try To Write The Standard
- The Purpose Of The Study Group Is To Try To Come Up With Objectives, Project Authorization Request, And Five Criteria To Get Approved To Become A Task Force – The Task Force Would Write A New Standard Or Modify An Existing One
- Why Might A New Standard Or Modifications Be Worthy of Study?
- The Market Is Demanding A Low-power, Low-cost, Low-latency Copper PHY

# Introduction (cont.)

- Target Applications For A Short-Haul PHY Are:
  - Server Rooms
  - Data Centers (Small To Medium Size)
  - Server Clusters
  - IP Based Storage Farms
  - Server Motherboards And Add-on Cards
  - Switch Aggregation
  - Blade Server Uplinks
  - Switch Stacking
  - ..But Not 10G To The Desktop
- There Is A Solution That Exists Today: CX4. But It Has Some Issues:
  - Twinax IB Cable Is Expensive
  - Can't Field Configure An IB Cable
  - Bend Radius Of IB Cable Is Large, Connector Is Large
    - Which Makes It Difficult To Work With In Tight Spaces
  - Can't Interoperate With Other PHYs
  - Reach Is A Bit Too Short: 15m

# Introduction (cont.)

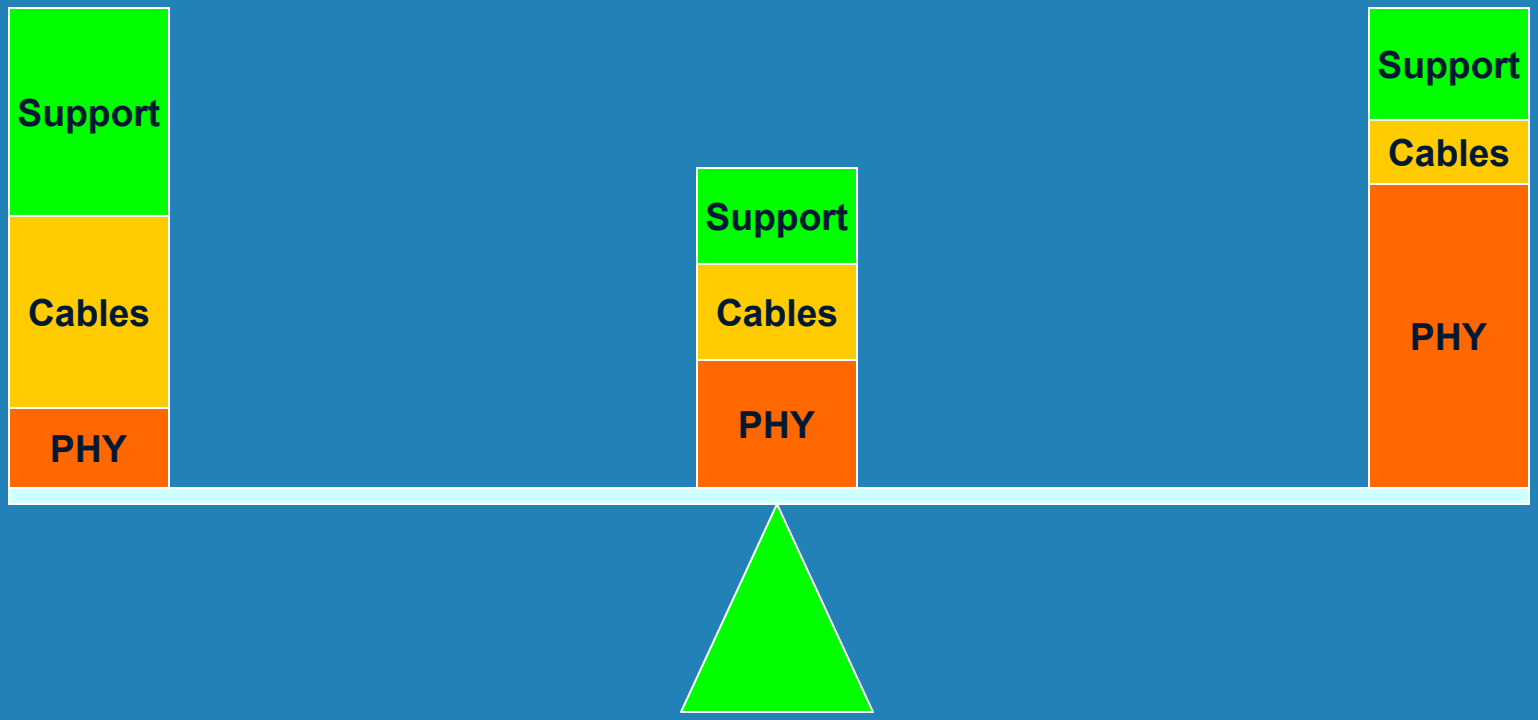
- There Is A Standard Which Addresses All Of These Issues:
  - 10GBASE-T
  - Runs On UTP Cable
    - Inexpensive
    - Field Configurable
    - Good Bend Radius, Flexible
    - Allows For Interop With 100BASE-T/1000BASE-T
    - Long Reach 100m (Cat6a/Cat7)
  - By Targeting 100m On UTP
    - The Power/Cost/Latency Is Driven By The Complexity Of Solving This Problem
    - The Resulting Solution Is Not Optimized For Short Reach Applications

# Introduction (cont.)

- Basic Point Of Short Haul Is To Target Shorter Reach
  - This Will Substantially Reduce Power/Cost/Latency
  - Shorter Reach Addresses The Target Applications
- Product Compatibility With Other Base-T PHYs will be considered
  - Auto-Negotiation
  - 100BASE-T/1000BASE-T/10GBASE-T
- Leverage Existing Standards As Much As Possible

# Total Cost of Ownership

Seeking Balance Between Cable And Silicon Cost/Complexity

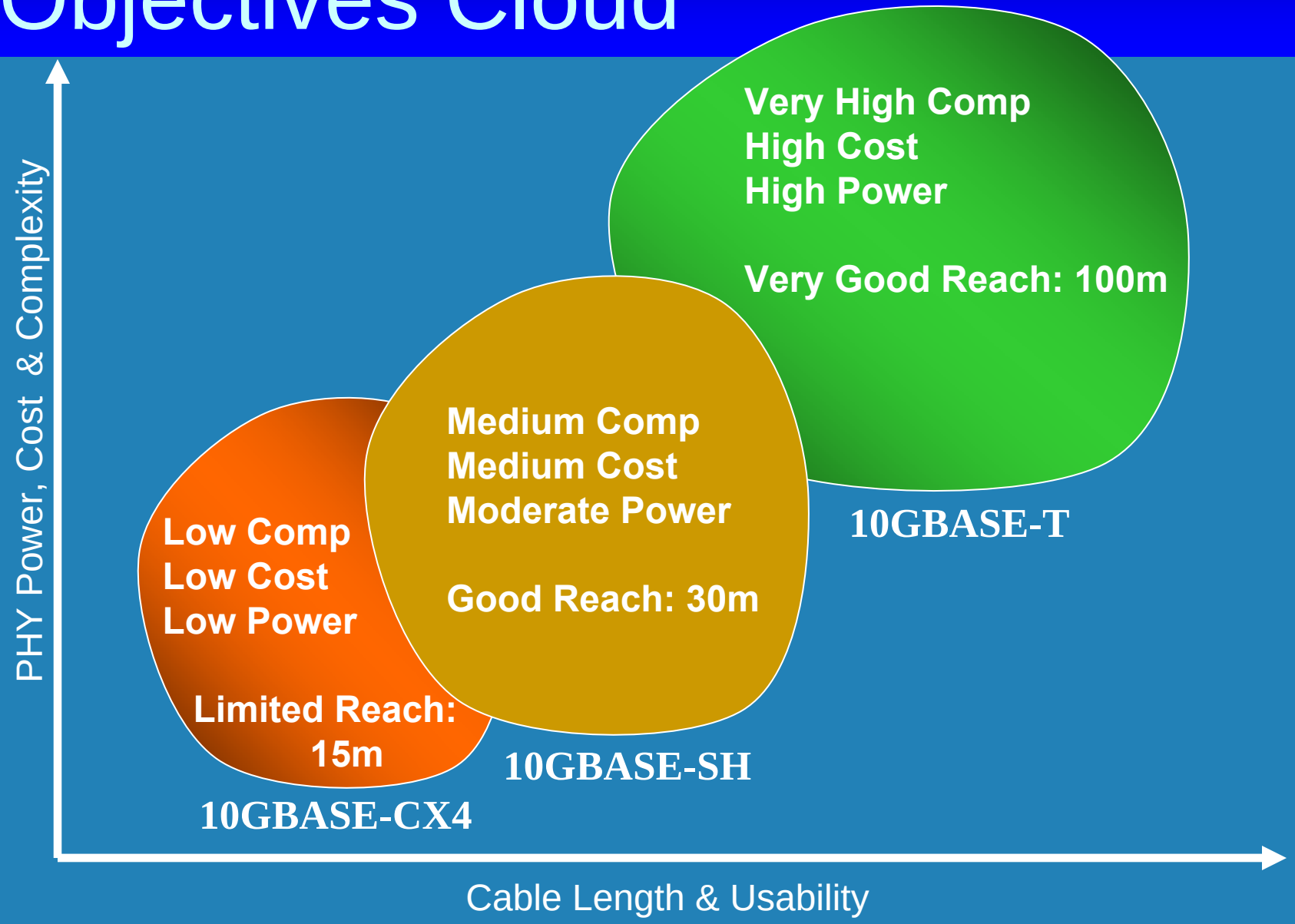



**10GB-CX4:**
Relatively Simple PHY
Complex & Costly Cables
Higher Support Cost to Maintain Custom Cables

**10GBASE-SH:**
Moderately Complex PHY
Low-Medium Complexity of Cables
Low Support Cost to Maintain Cable Infrastructure

**10GBASE-T:**
Very Complex PHY
Low Complexity of Cables
Moderate Support Cost to Maintain Cable Infrastructure (Mitigation)

# Objectives Cloud

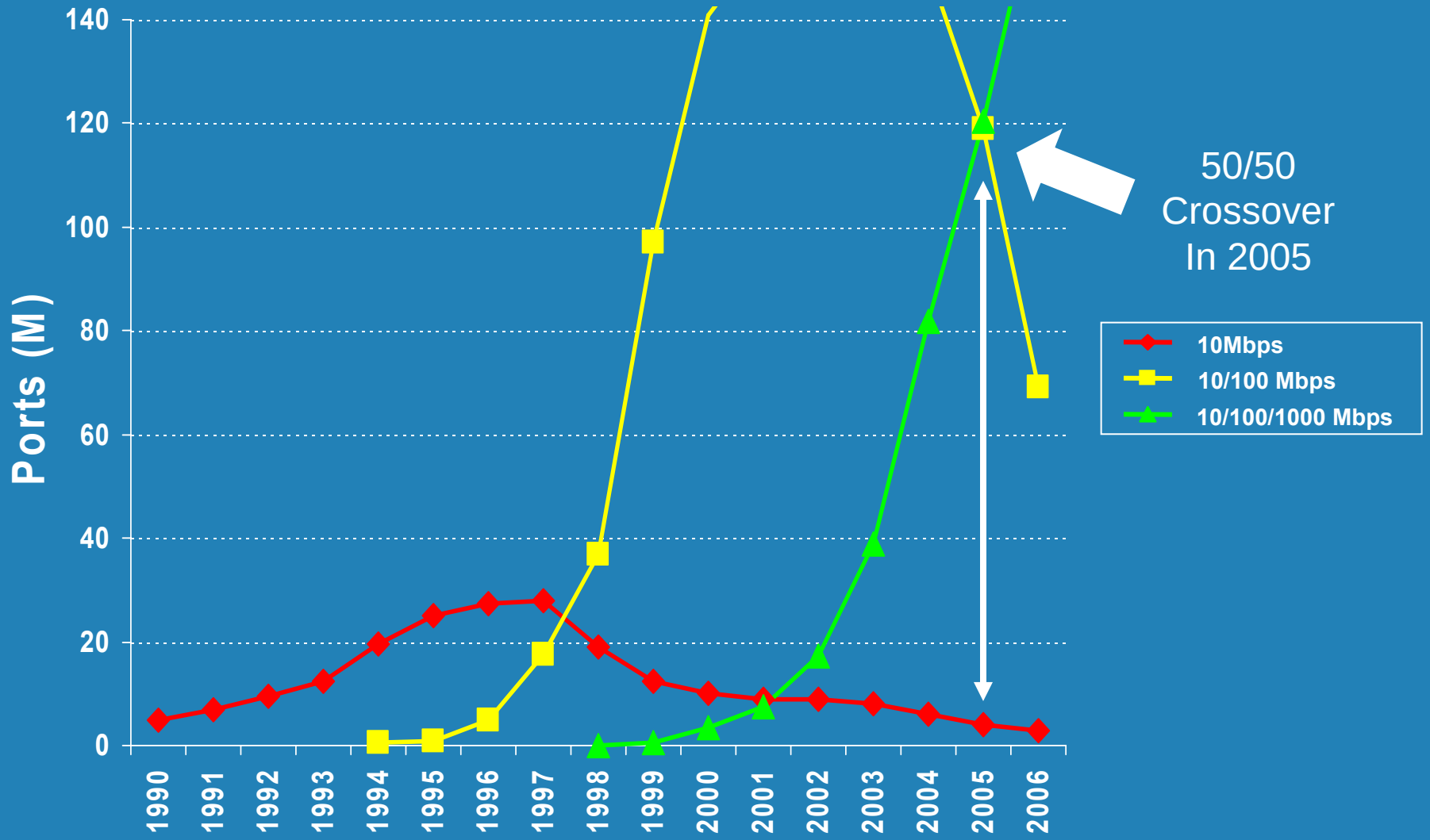


# Market Requirements and Potential

# Outline

- Market Requirements And Potential
  - Market  Will Be As Big Or Bigger Than 10GBASE-CX4
  - What Is The Market Demanding: Low Power, Cost, Latency
    - Field Configurable/Testable Cables
    - Be Able To Fit In XFP/XENPAK/X2
    - Be Able To Fit On PCI/PCI-E Card
    - Be Able To Put On Motherboard
  - Applications (Stacks, Aggregation, Server Clusters, Storage Arrays)
  - With GigE To The Desktop Reaching Over 50% Penetration And Ramping, The Market Is Demanding A 10G Pipe For Stacking/Aggregation That Can Auto Negotiate Multiple Speeds
  - If We Can Hit The Power/Cost/Latency Goals And Backwards Compatible With 100BASE-T/1000BASE-T, It Would Go On Every Server MB Shipped And Fill The Gap Between 10GBASE-CX4 And 10GBASE-T For Stackable Switch Uplinks.

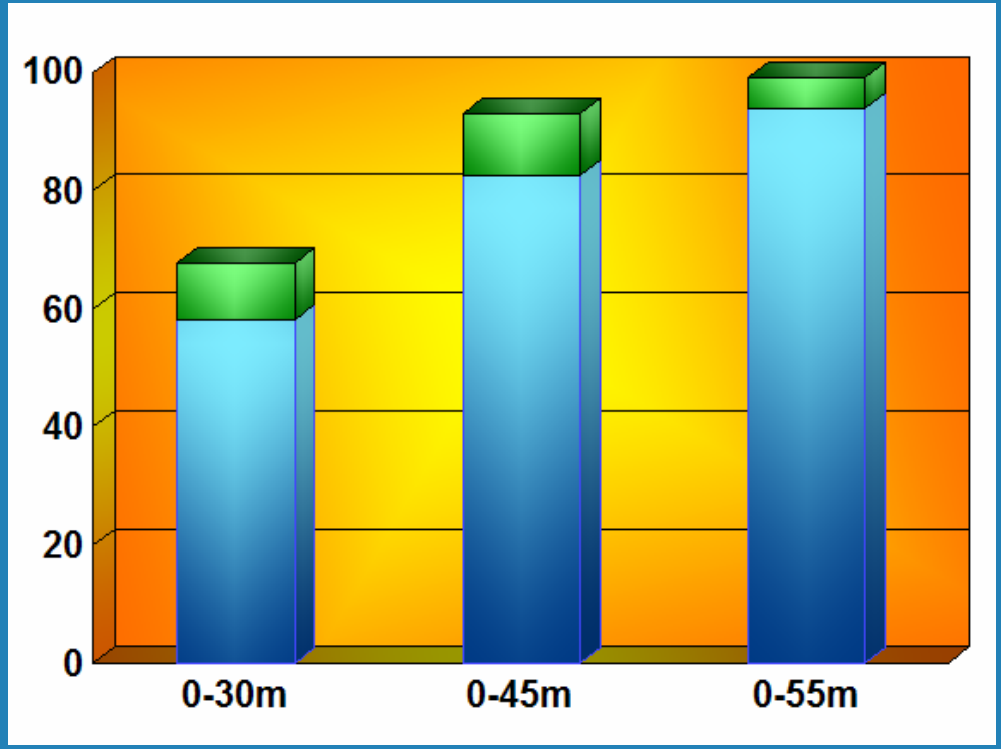
# Ethernet Technology Transitions




Sources: IDC, Cahners In-Stat, Cicada

# Data Center Connectivity

## Installed Cable Reach In Data Centers



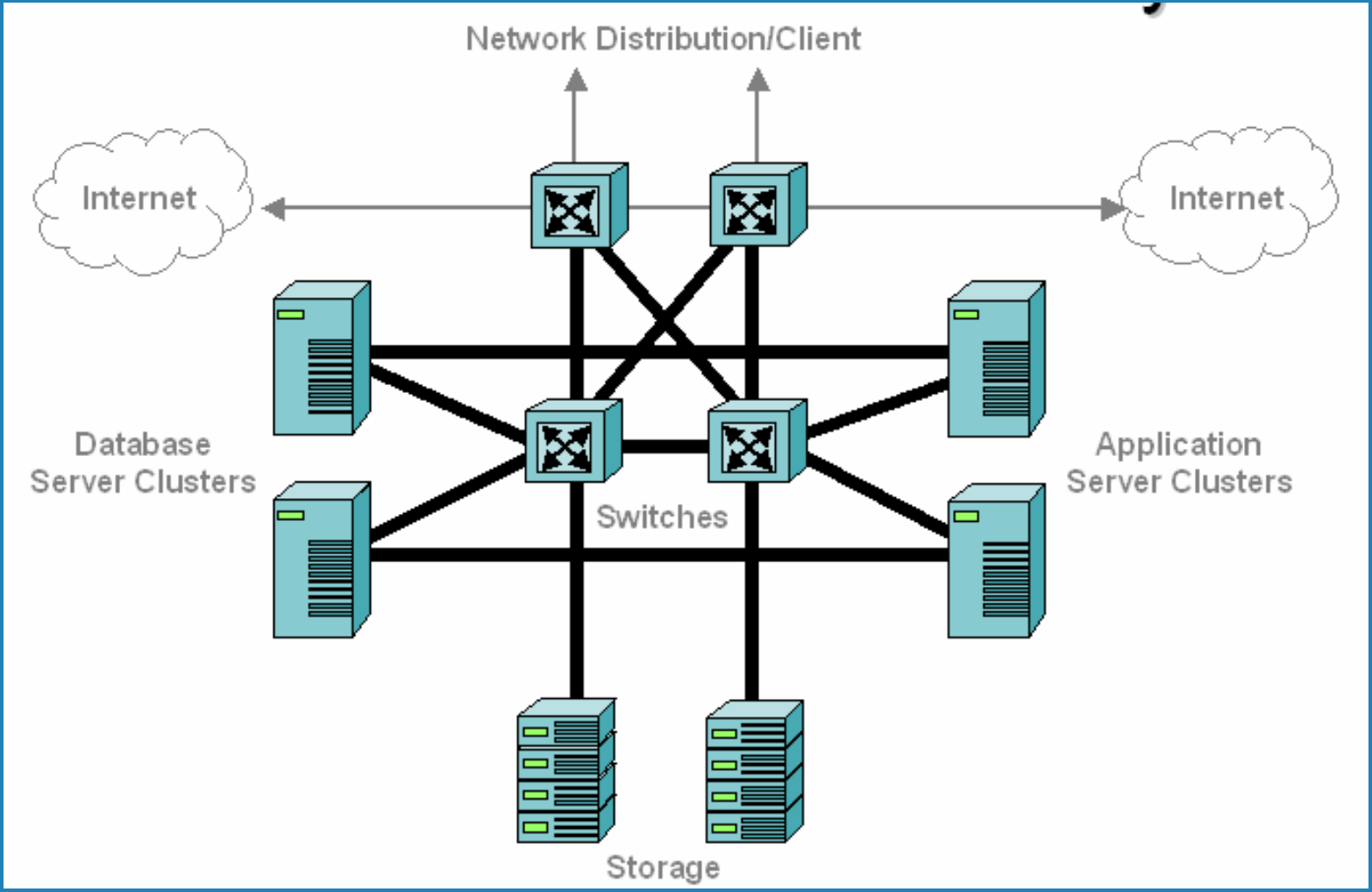

**Data From Installation Contractors And Based on 5000+ ft² (Big) Data Centers**

**90% Of All Data Centers Are ≤ 1000 ft² (roughly 5m x 19m)**

30m Covers 100% Of "Typical" Data Centers And 67% Of "Big" Data Centers!

Source:DiMinico_1_1103 802.3an SG

# Data Center Connectivity

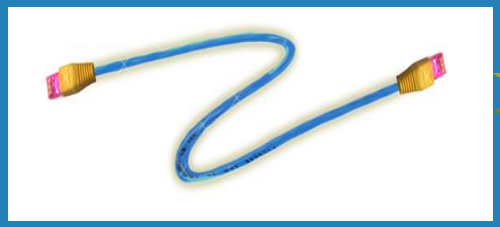


# Target Applications

Blade Servers
And Storage Arrays

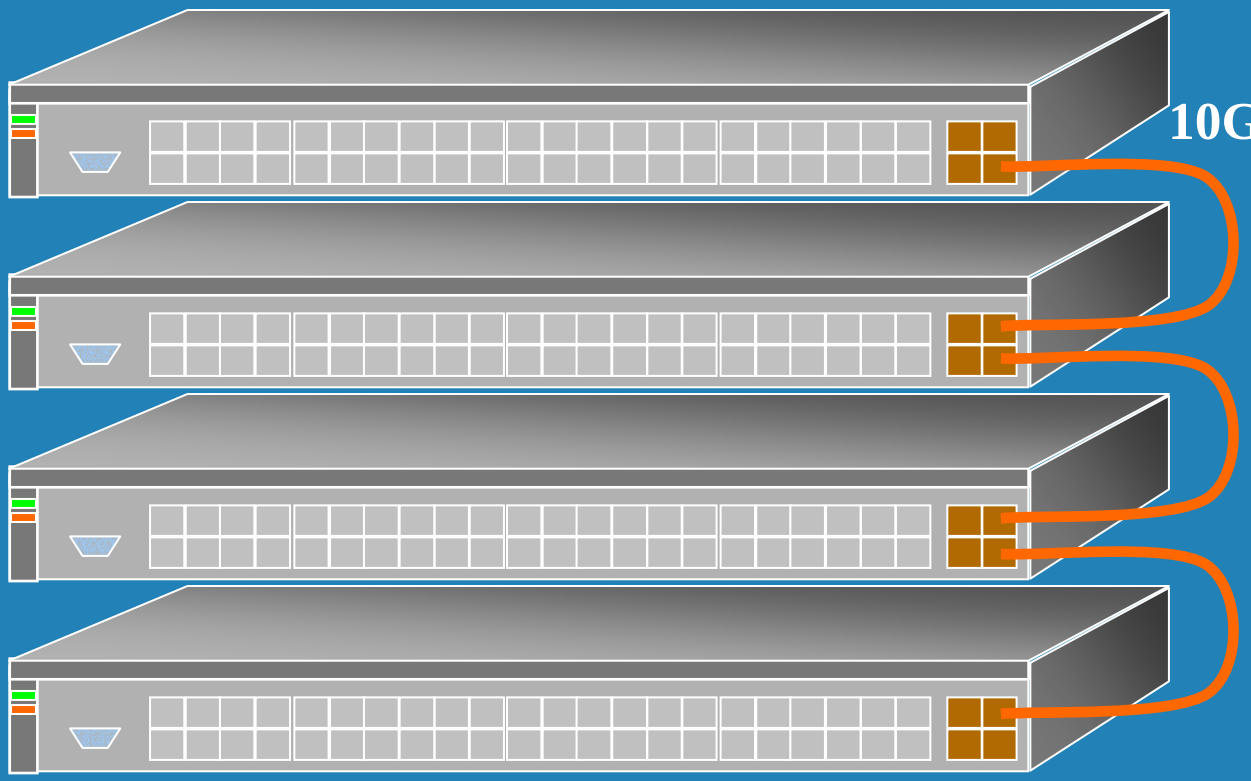
Copper
Modules

Inter-System
Links



Inter-System
Connectivity

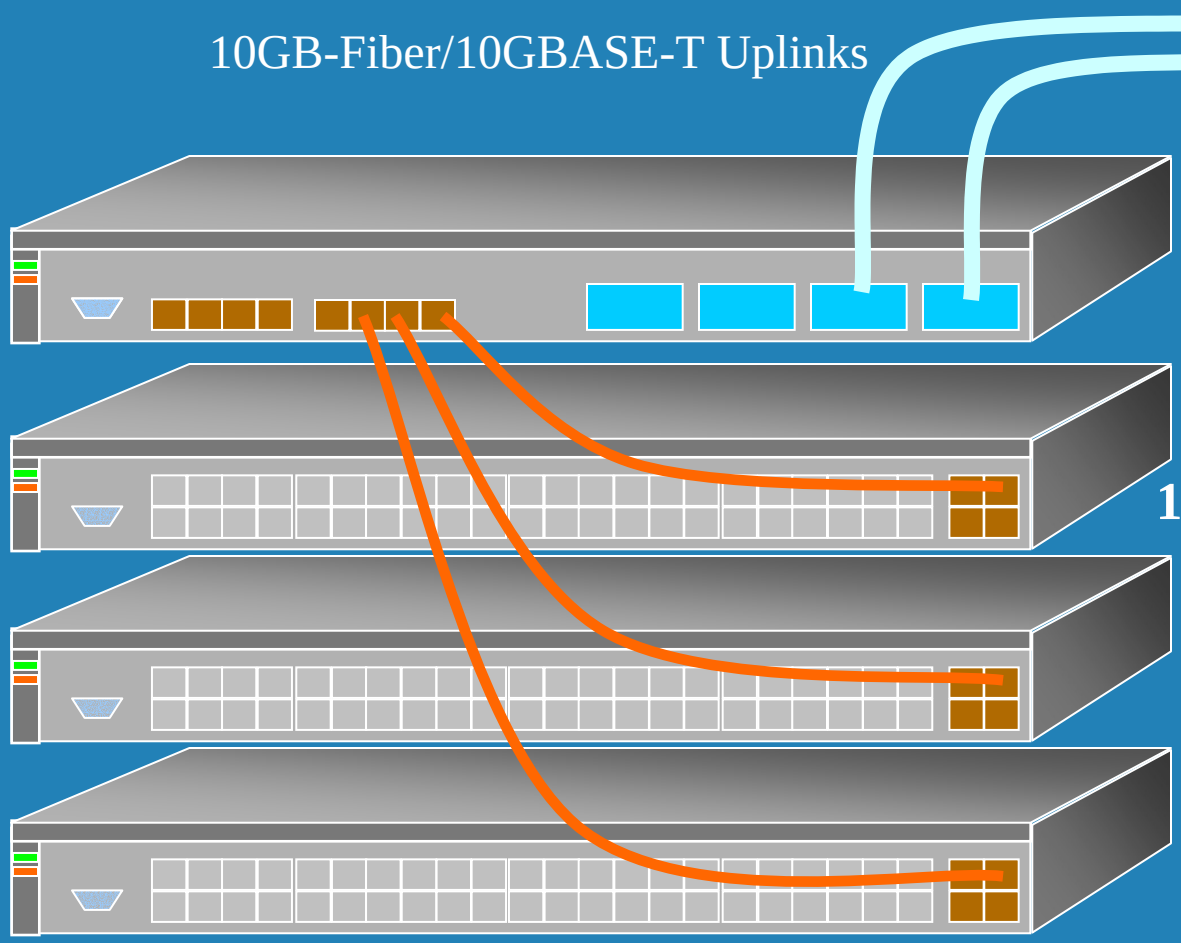
# Applications - Stackables



**10GBASE-SH Links**

**100/1000BASE-T Stackables**

# Applications - Aggregation

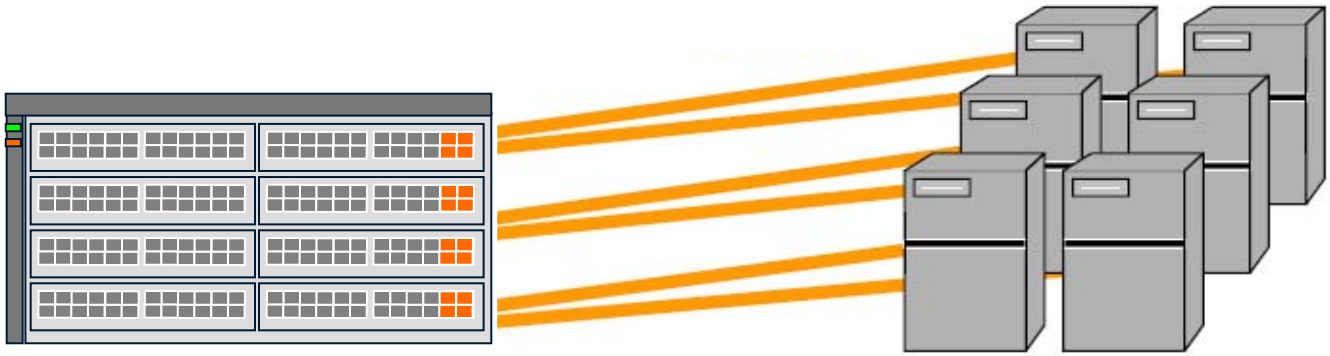


# Applications – Server Clustering

**Local Switching**

**10GBASE-SH Links**

**Servers**



# Cabling

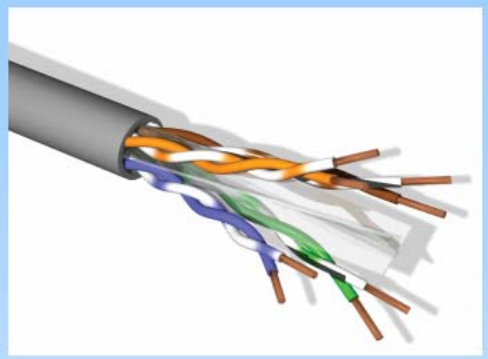
# Cabling

## Outline

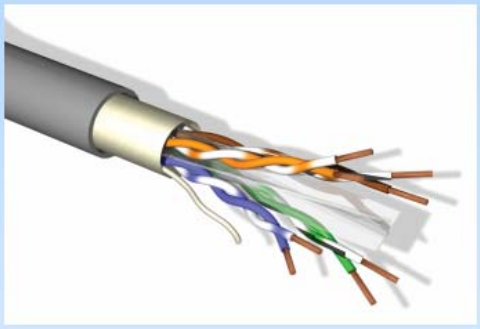
1. Brief Review Of Cable & Connector Options
2. Performance Advantages Of Screened Cabling
3. Screen Management & Installation Practices
4. Size/Flexibility/Cost/Installation Aspects
5. Deployment/Experience With Screened Cabling

# Cat 6 Cable & Connectors

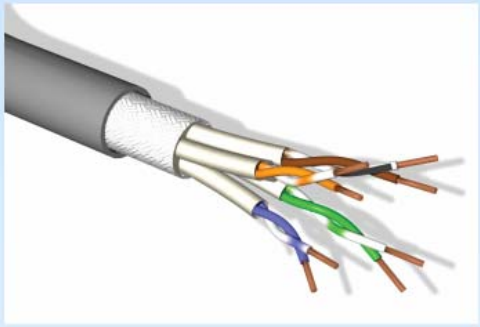
UTP Cable



FTP Cable

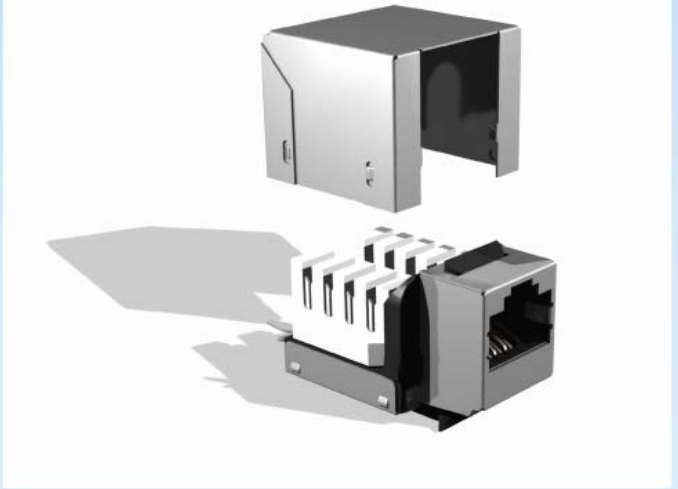


STP Cable



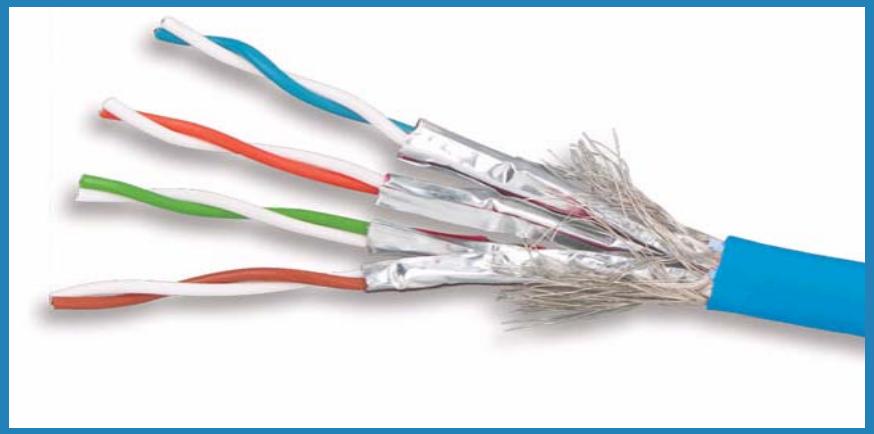
unscreened RJ-45 connector

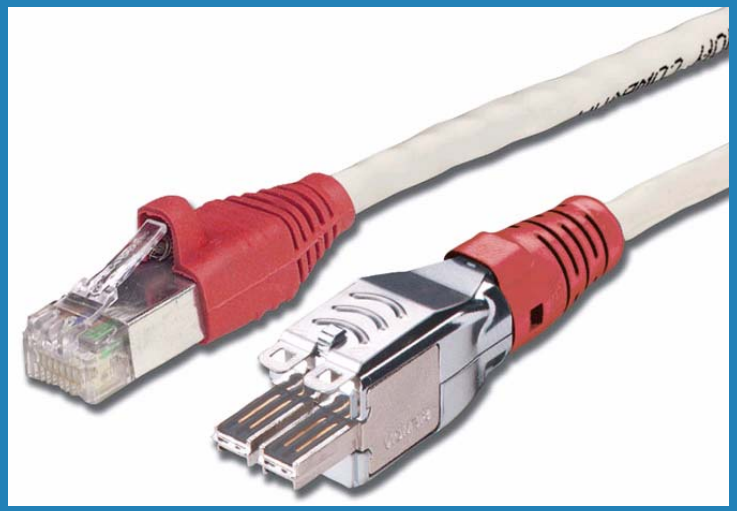
screened RJ-45 connector



# Cat 7 Cable & Connectors

**Category 7 cable (STP)**



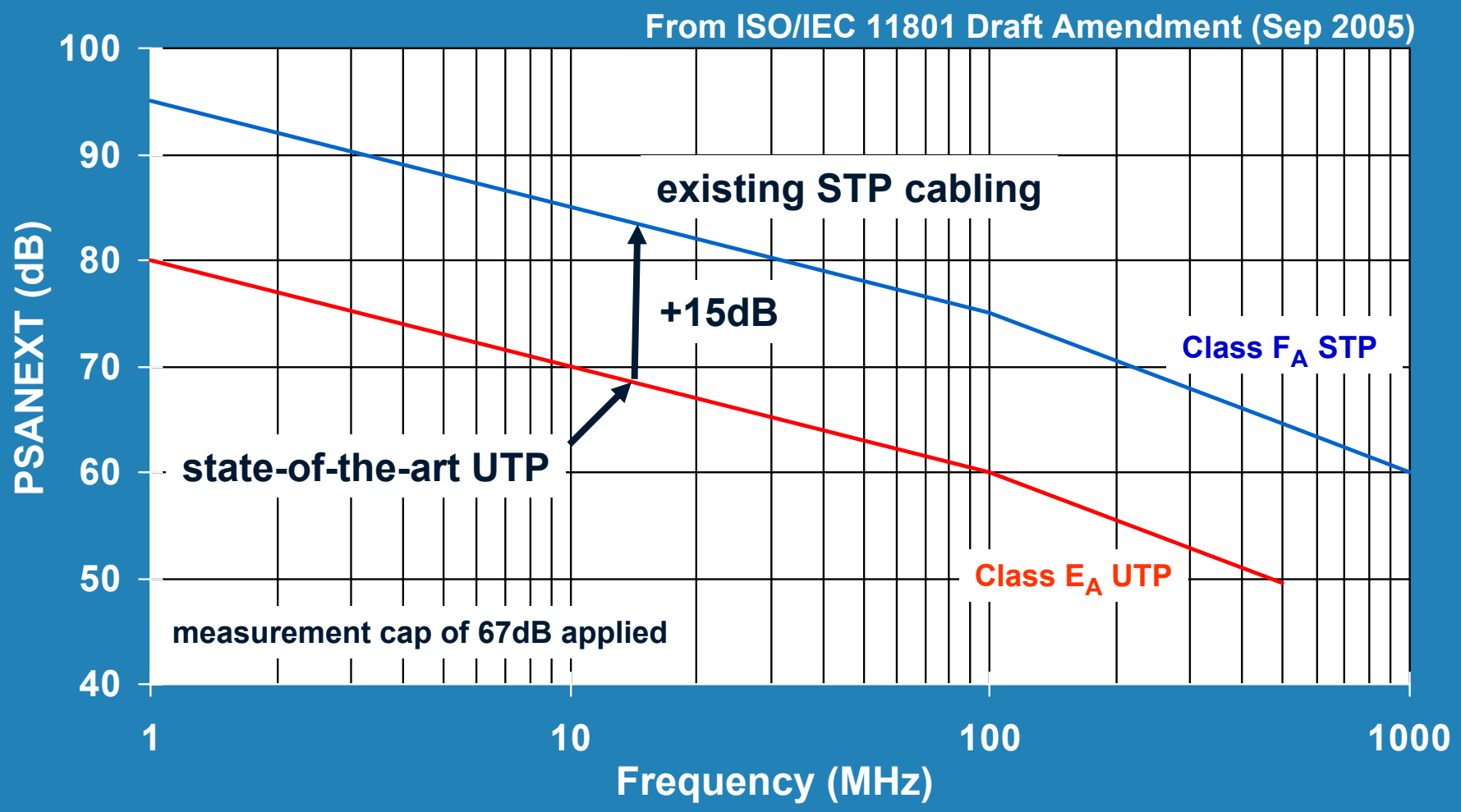


**Cat 7 connectors specified by ISO/IEC 11801 cabling standard:**

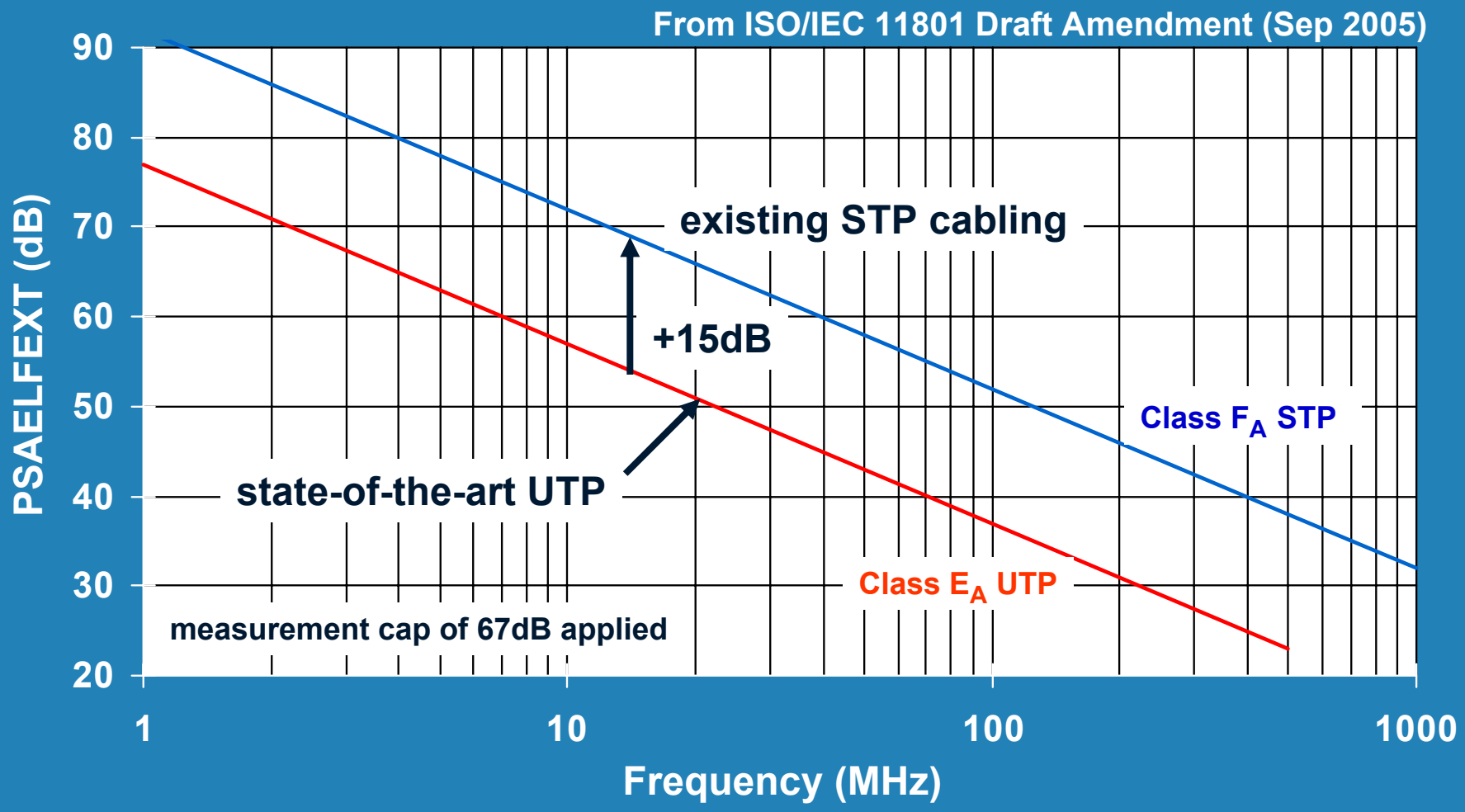
**IEC 60603-7-7 (RJ-45 type)**

**IEC 61076-3-104**

# Channel PSANEXT (100m cable with 4 connectors)

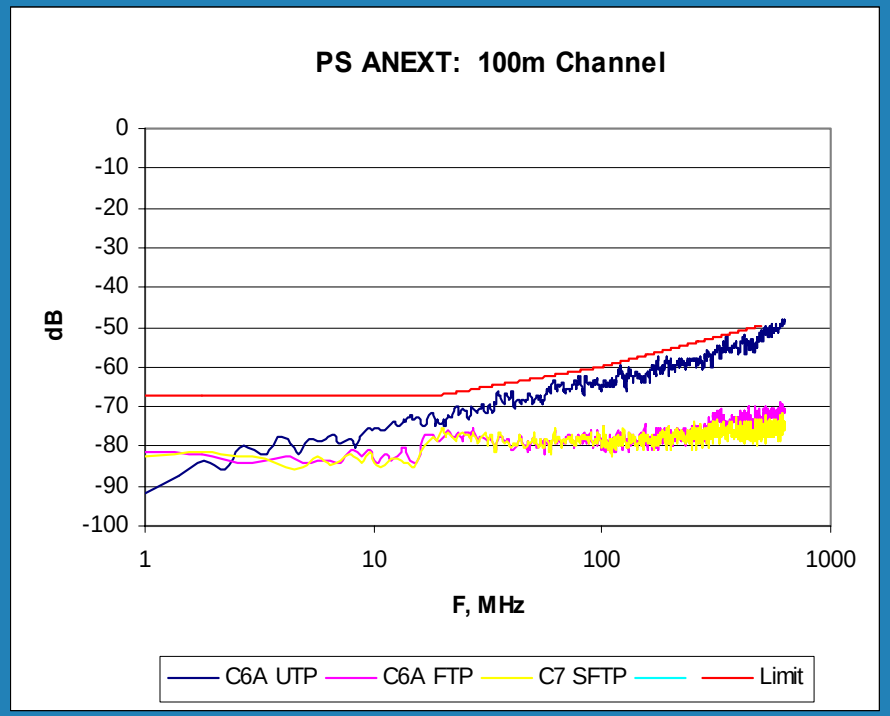


# Channel PSAELFEXT (100m cable with 4 connectors)

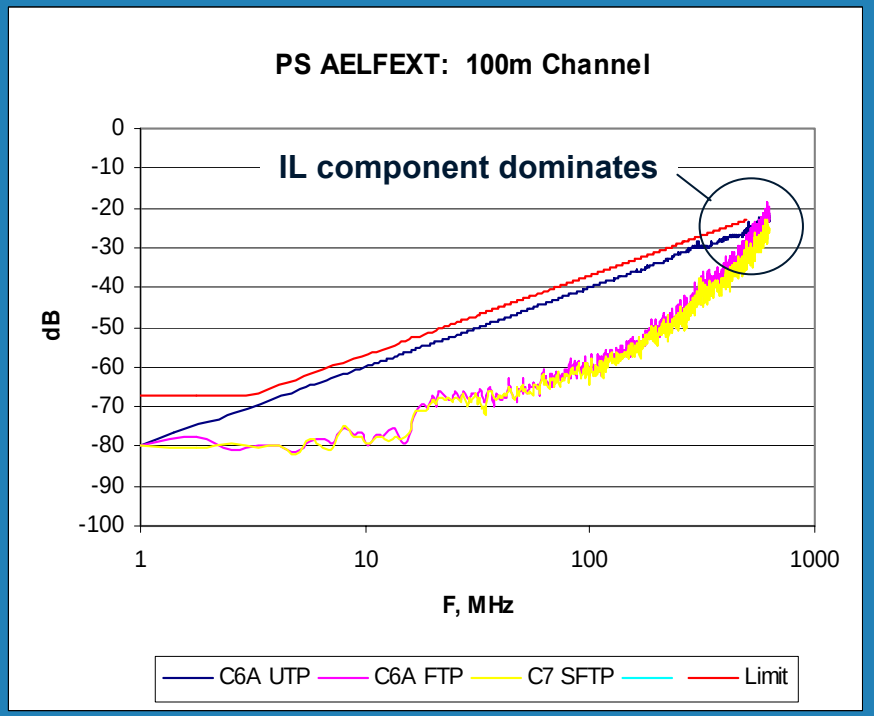


# Alien Crosstalk for 3 connector 100m Channels

source: Siemon Company



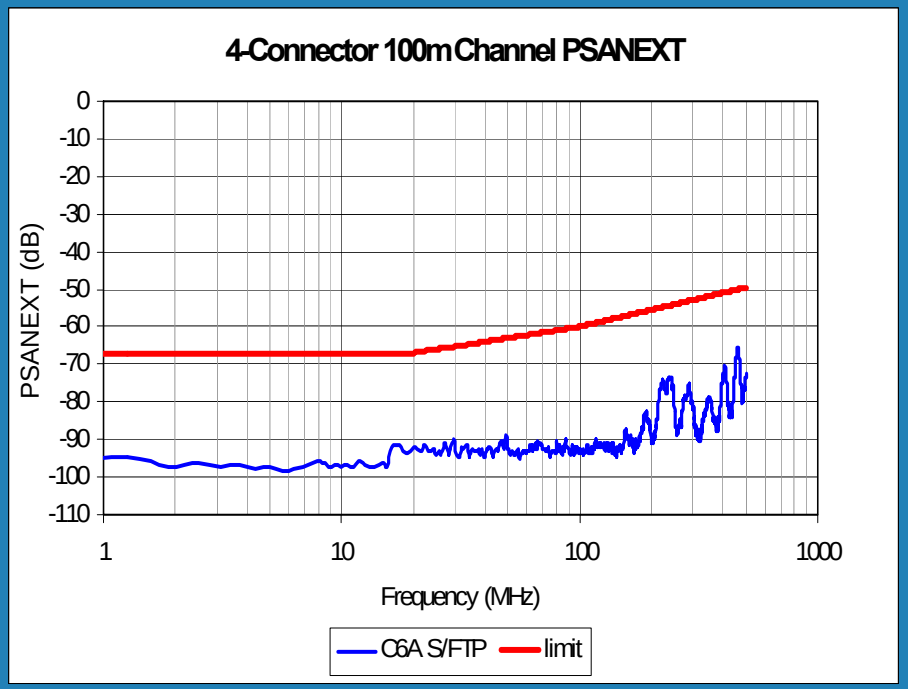

source: Siemon Company



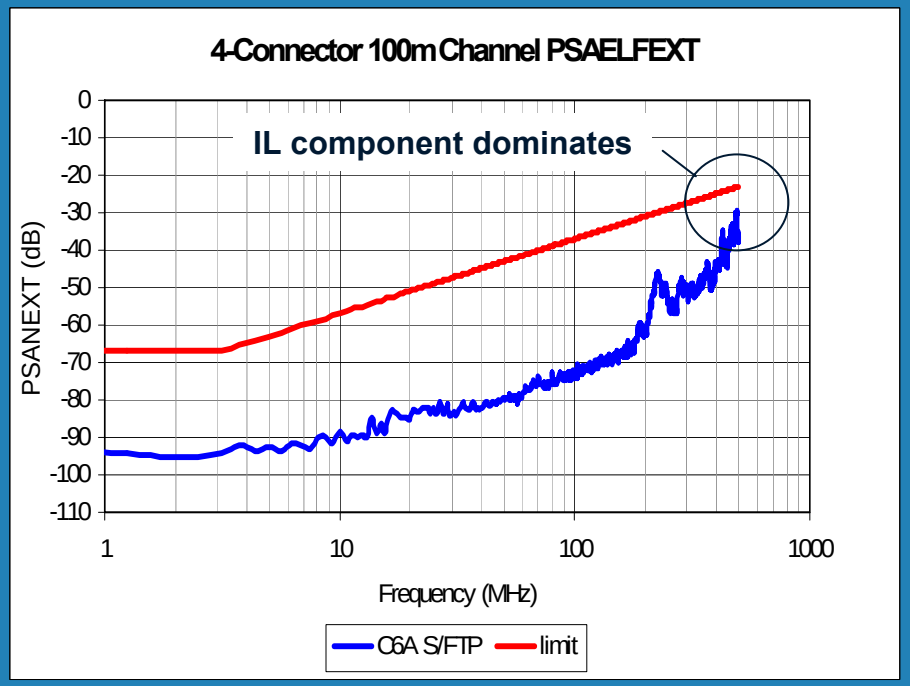

1. Data Using "Six Around One" Alien Crosstalk Measurement Setup
2. C6A = Cat 6 Augmented Or Class $E_A$ Cabling, C7 = Cat 7 Or Class $F_A$ Cabling
3. Class $E_A$ Limits Specified By ISO/IEC 11801 Draft Amendment (Sep 2005)

# Alien Crosstalk for 4 connector 100m Cat $6_A$ STP Channel

source: Brand-Rex



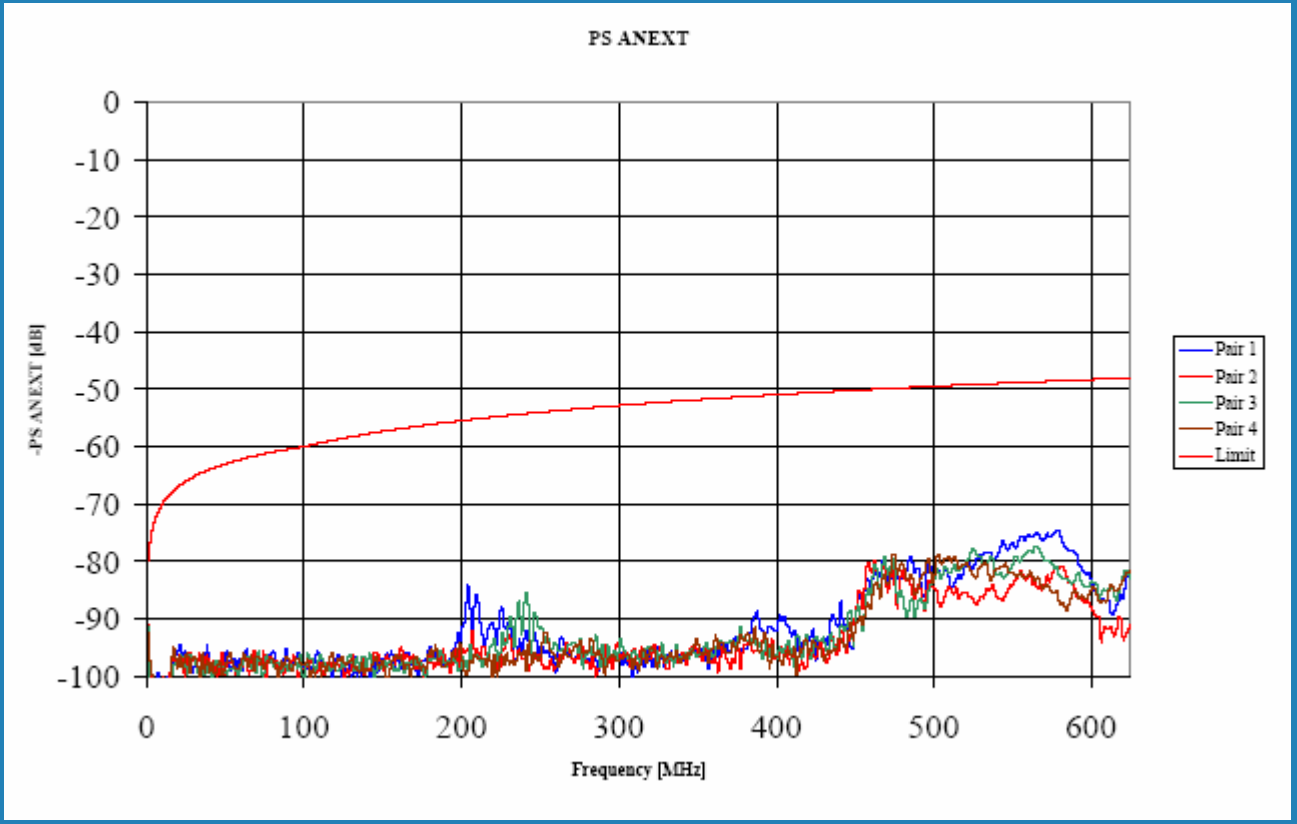

source: Brand-Rex




1. **Data Using "IEC Drum Method" Alien Crosstalk Measurement Setup**
2. **C6A = Cat 6 Augmented Or Class $E_A$ Cabling**
3. **Class $E_A$ Limits Specified By ISO/IEC 11801 Draft Amendment (Sep 2005)**

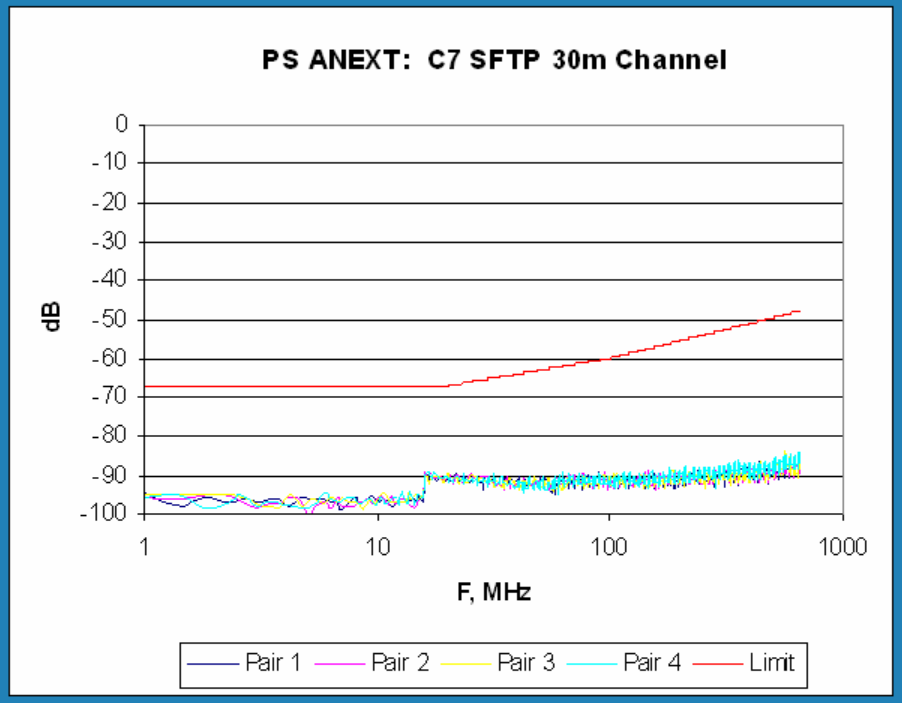
# Alien Crosstalk for 4 connector 100m Cat 6 STP Channel

**source: Tyco Electronics**

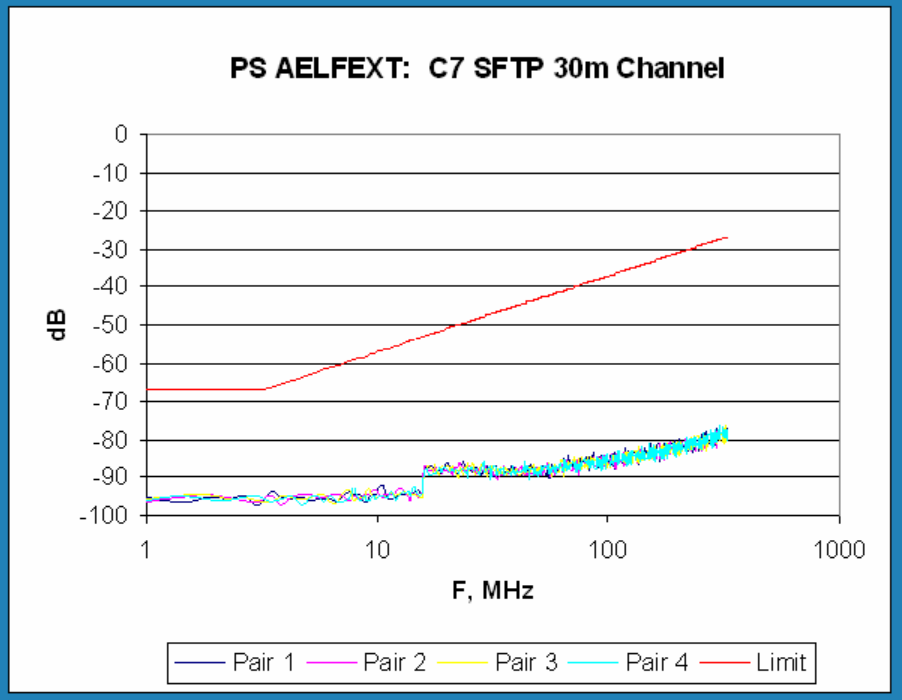


# Alien Crosstalk for 3 connector 30m Cat 7 STP Channel

source: Siemon Company



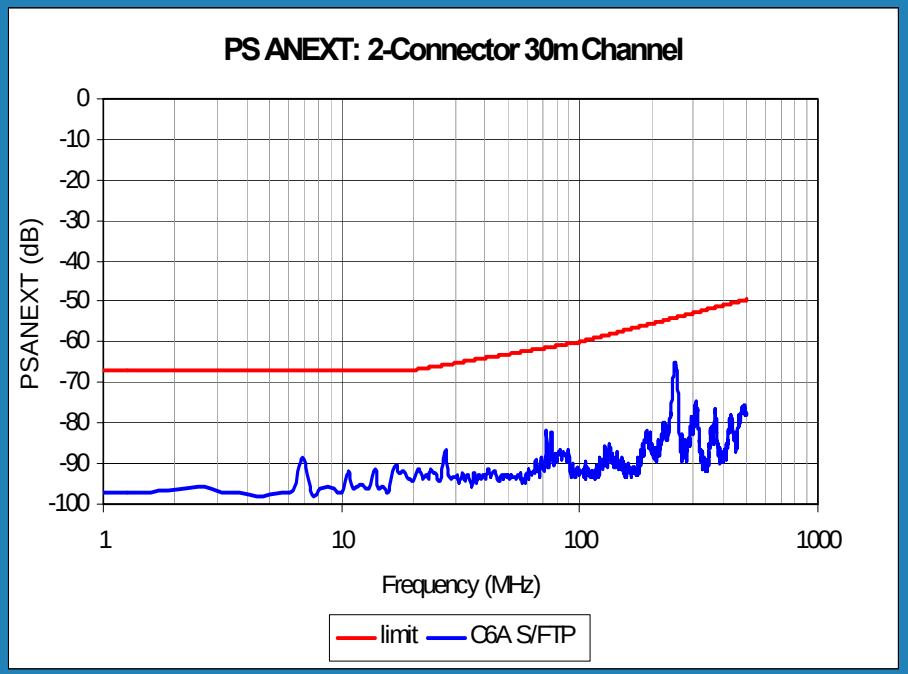

source: Siemon Company



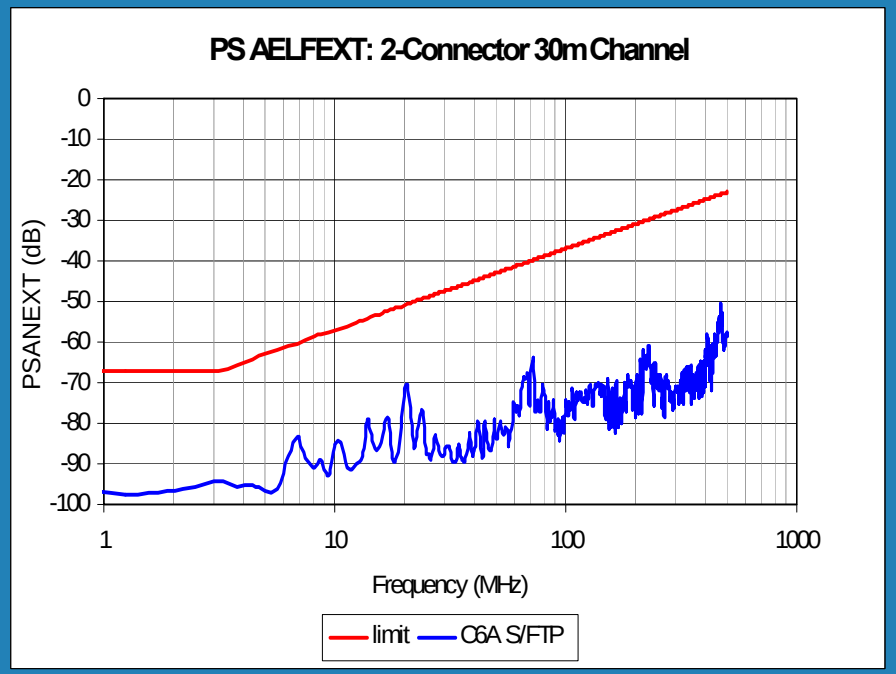

1. data using "six around one" alien crosstalk measurement setup
2. Class $E_A$ limits specified by ISO/IEC 11801 draft Amendment (Sep 2005)

# Alien Crosstalk for 2 connector 30m Cat $6_A$ Channel

source: Brand-Rex




source: Brand-Rex




1. **Data Using "IEC Drum Method" Alien Crosstalk Measurement Setup**
2. **Class $E_A$ Limits Specified By ISO/IEC 11801 Draft Amendment (Sep 2005)**

# Screened vs Unscreened Channel Performance (100m cable with 4 connectors)

Based On ISO/IEC 11801 Draft Amendment & Draft TR-24750 (Sep 2005)

| Cable Type → / Parameter ↓ | Class E UTP | Class $E_A$ UTP | Class E FTP | Class E STP | Class F STP | Class $F_A$ STP |
|---|---|---|---|---|---|---|
| Max Freq (MHz) | 500 | 500 | 500 | 500 | 600 | 1000 |
| Insertion loss (dB @ 500MHz) | 53.4 | 49.3 | 53.4 | 53.4 | 49.3 | 46.7 |
| Return Loss (dB @ 500MHz) | 6.0 | 8.0 | 6.0 | 6.0 | 8.0 | 8.0 |
| NEXT (dB @ 500MHz) | 22.0 | 27.9 | 22.0 | ~30.0 | 52.4 | 52.4 |
| ELFEXT (dB @ 500MHz) | 9.3 | 9.3 | 9.3 | ~25.0 | 32.6 | 32.6 |
| PSANEXT (dB @ 500MHz) | - | 49.5 | ~57.0 | ~64.5 | 64.5 | 64.5 |
| PSAELFEXT (dB @ 500MHz) | - | 23.0 | ~30.0 | ~38.0 | 38.0 | 38.0 |

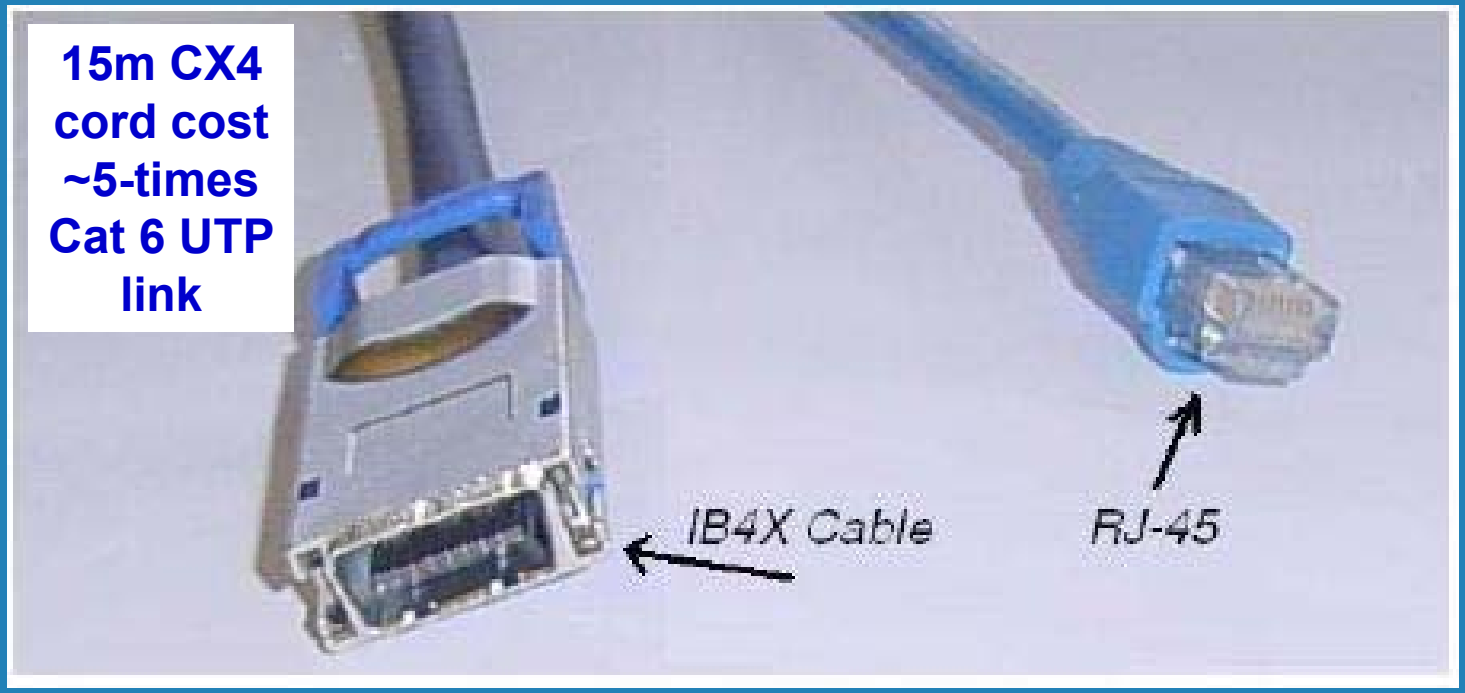
biggest incremental gains

# Comparison of Cabling Types

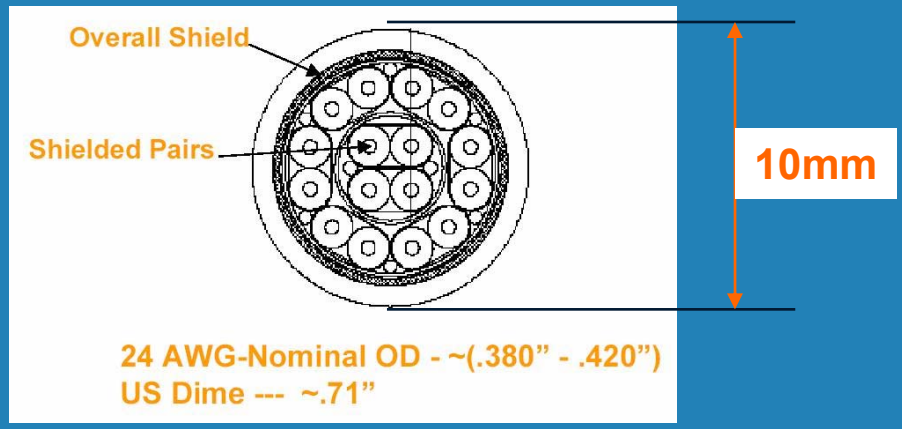
| Cable Type Parameter | Cat 6 UTP | Cat $6_A$ UTP | Cat 6 FTP | Cat 6 STP | Cat $7/7_A$ STP | optical fibre |
|---|---|---|---|---|---|---|
| Cable dia (mm) | 6.2 | ≤ 9.0 | 7.0 | 7.2 | 7.5 | 4.6 |
| Min bend radius when pulling (mm) | 49.6 | ≤ 72.0 | 56.0 | 57.6 | 60.0 | 70.0 |
| Min bend radius installed (mm) | 24.8 | ≤ 36.0 | 28.0 | 28.8 | 30.0 | 45.0 |
| cost per cable metre (factor x Cat 6 UTP) | 1.0x | 1.8x | 1.45x | 1.5x | 1.6x | 2.5x |
| component costs 50m channel (fxC6U) | 1.0x | 1.8x | 1.25x | 1.4x | 1.5x | 3.0x |

Notes: specified by TIA/EIA-568-B.2 Addendum 11
8x cable OD, according to IEC cable specifications
4x cable OD, according to IEC cable specifications
average 2005 manufacturer selling prices for PVC cables
based on average 2005 MSPs for components only, including 4 connectors

# 10GBASE-CX4 Cable & Connector







# Screen Management

- Huge Amount Of Screened Cabling Deployed In Europe
  - Germany 98.4%, France 89.3%, Austria 87.3% (BSRIA 2004)
- At Least 30 Years Experience In Screen Management Has Been Put In Public Domain Practices & Guides
  - E.G. CENELEC EN 50174 Cabling Planning & Installation
- Contribution To Installation Process Is Estimated To Be Small (Additional 10-20% Labour Time/Cost)

# Why Screened Cabling for 10 Gig Short Haul Copper?

1. At Least 15 dB Advantage In Alien Crosstalk Over State-of-art UTP
2. Cable Thinner & More Bendable Than 10GBASE-CX4 Cable
3. Reduced Segregation Distance With Power Cables
4. Higher Patching Density Possible
5. Installation & Management Aspects Well Understood
6. Reduced RF Emission & Increased Noise Immunity
7. Costs No More Than State-of-art UTP
8. Costs Much Less Than Optical Fibre

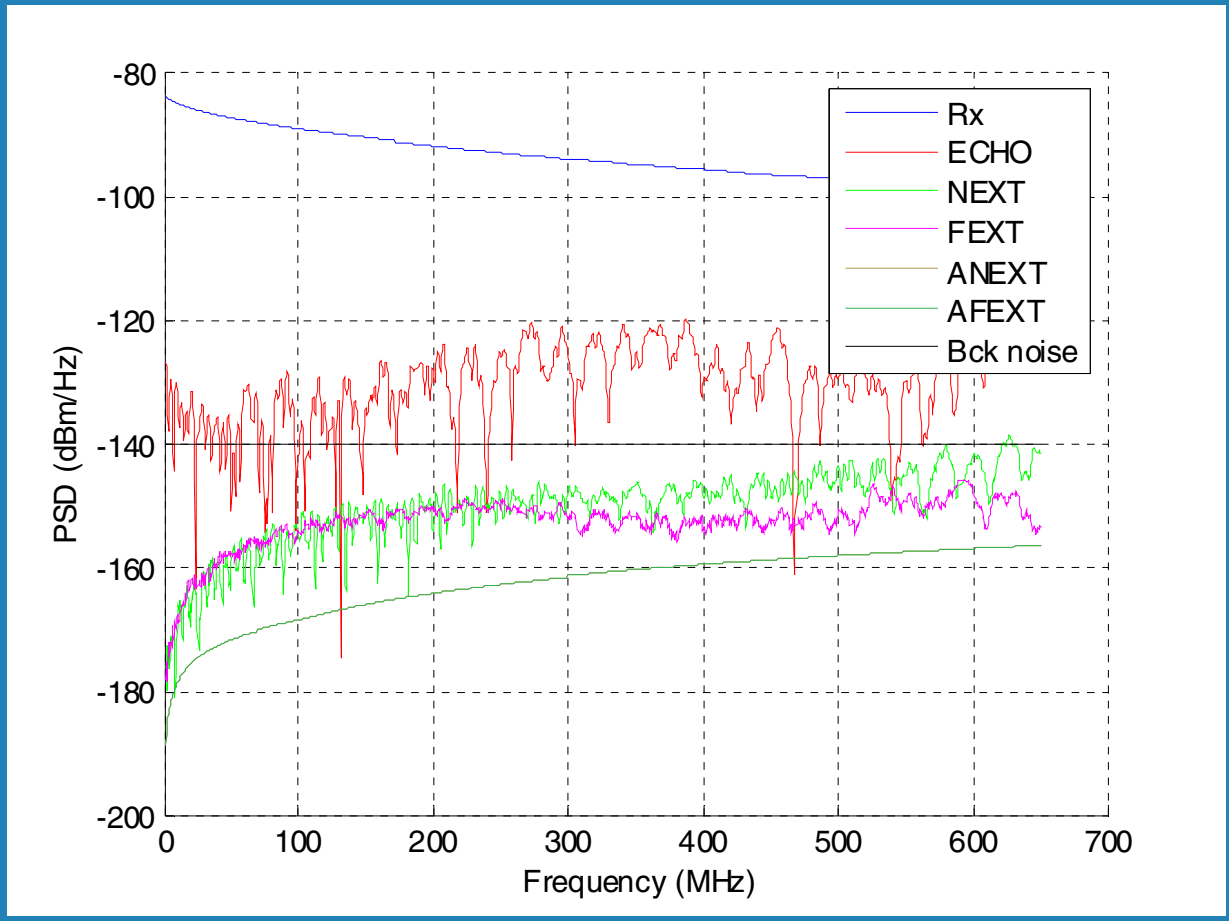
# Technical Feasibility

# Meeting The Power Goal

- Analysis Of 30m 3 Link Segment Cat 7 Channel Models Indicates That Significant Power Reduction Is Possible
  - No Need For FEXT Cancellation
  - Very Small Or No NEXT Cancellation
  - Reduced Echo Cancellation Requirements
  - No Need For FEC
  - Significantly Reduced ADC Requirements

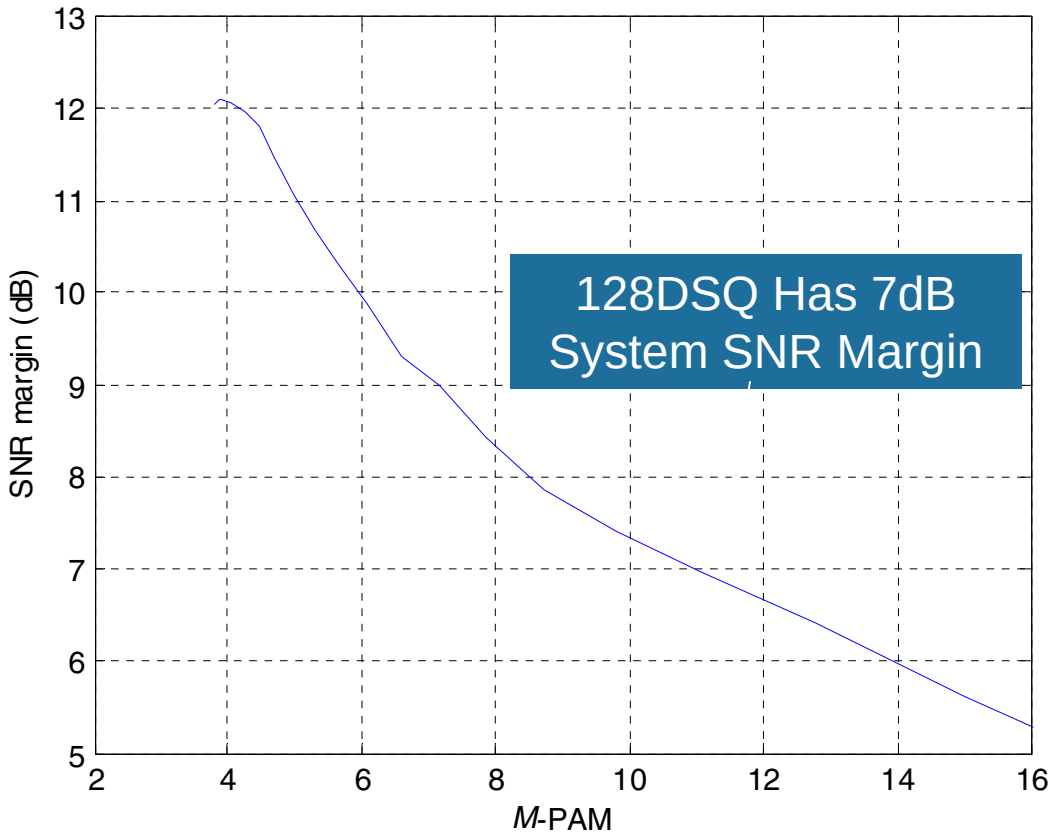
AND It Is Possible To Maintain Significant System SNR Margin!

# Relative PSD




Channel: 30m Cat 7, 3 Link Segments
-140 dBm/Hz Background Noise
Echo Is The Dominant Impairment

# System SNR Margin




30m Cat-7, 3-Link-Segments

-140 dBm/Hz Background Noise

Minimal Echo Cancellation

No NEXT/FEXT Cancellation

No FEC

# Meeting The Latency Goal

- Removing The LDPC Coding Significantly Reduces Latency
- Shorter Reach Significantly Reduces Interference Cancellation Processing Requirements Which Further Reduces Latency

# Cost/Complexity/Power

- By Targeting 30m Reach On Cat7 Cable DSP Complexity Is Reduced By At Least 70%
  - Reduces Gate Count Proportionately
- AFE Performance Requirements Are Reduced By At Least 12dB
  - Leads To Further Area And Power Reductions
- Overall Power/Cost Is Reduced By 75%

# Conclusion

# Conclusion

- We Need A Short Haul 10G Copper PHY
  - To Improve The Usability Of 10G Short Haul Interconnect
  - To Lower The Power, Cost, Latency For High Density/Performance Applications
  - Meets The XFP/X2/XENPAK Module Size & Power Envelope
  - To Break Open The Bottleneck Of Aggregation And Stacking For 1G Switches
  - To Enable High Density Server Clustering
  - To Support The Sweet Spot Of PHY And Media Costs/Usability

# Conclusion

- The PHY Would Have Broad Market Potential
  - Every Server Is A Potential Socket
  - Definitely Larger Than The Market For 10GBASE-CX4
- The PHY Is Technically Feasible
  - Showed Plenty Of Margin At 30m With No Coding

# Conclusion

- Short Reach PHYs That Meet Most Of These Goals Are Being Brought To Market Today By Multiple Vendors
  - These PHYs Are Not Interoperable
  - Nor Do They Interoperate With 10GBASE-T
- End Users, OEMs, And Vendors Will All Benefit From A Standardized Solution That Will Guarantee Interoperability

**It's Time To Begin a Study Group Now!**

# Individual Supporters

- Allan Nielsen, Tyco Electronics
- Alan Flatman, LAN Technologies
- Alan Kwentus, KeyEye
- Brad Booth, Quake Technologies
- Chuck Hudson, HP Server Networking
- Colin Mick, Mick Group
- Dan Dove, DNS/HP ProCurve
- David Koenen, HP Server Networking
- Eric Lawrence, Nexans
- Gaurav Malhotra, Nexans
- Geoff Thompson, Nortel ?
- Harvey Scull, KeyEye
- Iain Ballingall, Brand-Rex
- Jeffrey Lynch, IBM
- Jim Barnette, Vitesse
- Jim Heckroth, Silicon Labs
- Joel Goergen, Force10
- John Siemon, Siemon
- Joseph Babanezhad, Plato Networks
- Keith Conroy, AMCC
- Ken Hodge, Brand-Rex
- Mike Bennett, Lawrence Berkley
- Mike McConnell, KeyEye
- Nick van Bavel, Vitesse
- Olindo Savi, Siemon
- Phil Callahan, Silicon Labs
- Sudeep Bhoja, Independent
- Susumu Hara, Silicon Labs
- Terry Hewlett, NetEffect
- Valerie Rybinski, Siemon
- Wayne Mueller, NetEffect
- Wei Fu, AMCC
- Worayot Lertniphonphun, Vitesse
- Yakov Belopolsky, Bel Stewart
- Ziad Hatab, Vitesse

# Comment/Question Period

Comments and Questions?

# Straw Poll

61 Individuals Support the Formation of a
**Short Haul 10G PHY Study Group**

35 802.3 Voting Members Support the Formation of a
**Short Haul 10G PHY Study Group**

21 Companies Support the Formation of a
**Short Haul 10G PHY Study Group**

45 Individuals would attend and contribute to a
**Short Haul 10G PHY Study Group**