

Broad Market Potential and Economic Feasibility of the 100Gb Backplane and Cu Cable Solutions for the Volume Blade & Rack Server Markets

**IEEE 802.3 100Gb/s Backplane and
Copper Cable Study Group**

Singapore

March 2011

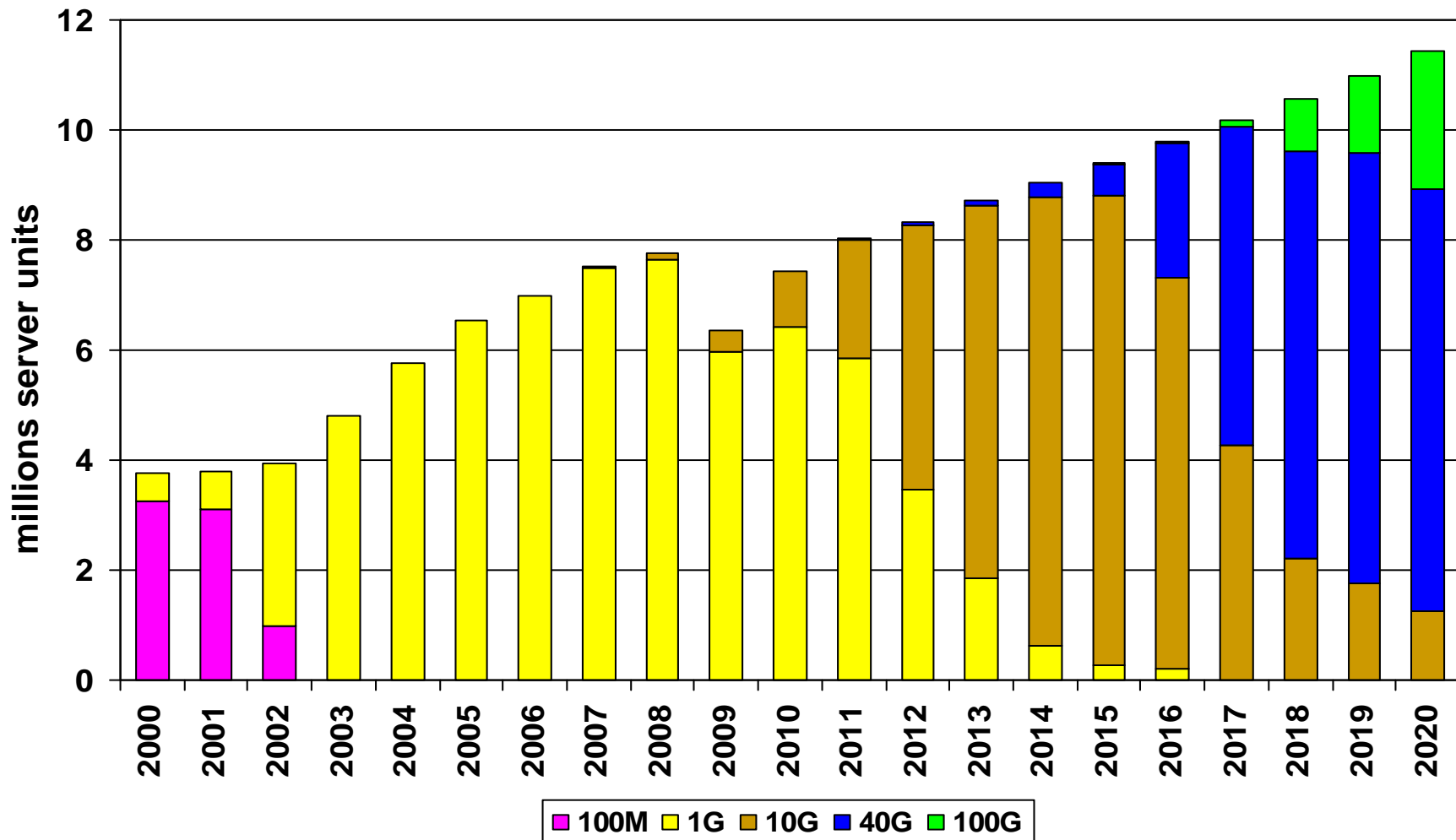
David Chalupsky, Intel Corporation

Summary Points for 10Gb in the Server Market

- Port mix: 100G will coexist with 10G & 40G.
- Smooth Adoption of 100G can be facilitated by:
 - **Balanced cost**
 - pay as you go
 - PHY vs Interconnect
 - 25Gb NRZ Channel is very challenging for the server market!
 - **Multi-speed ports**
 - decouples the NIC & Switch upgrade
 - Support Autonegotiation
 - **Plug compatibility**
 - speeds & media

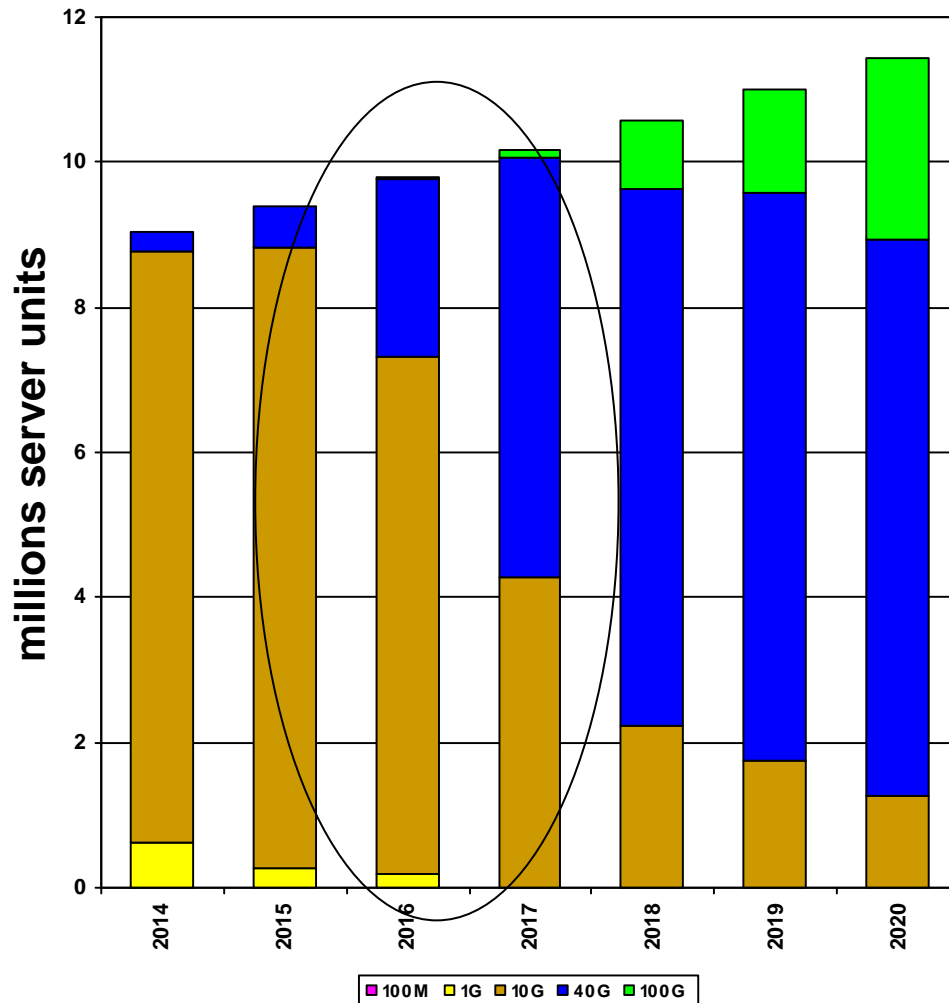
Revisit x86 Servers by Ethernet Connection Speed (2010 Forecast) From CFI

Based on IDC (2010) Server Forecast and hays_01_0407 ratios of Ethernet port speed



Port Mix at Introduction

Based on IDC (2010) Server Forecast and hays_01_0407 ratios of Ethernet port speed



At introduction, 100G server ports will coexist with 10G & 40G...
Even some 1G

Blade and Rack Servers should support all these speeds

Avoid putting a cost burden on 10G/40G ports

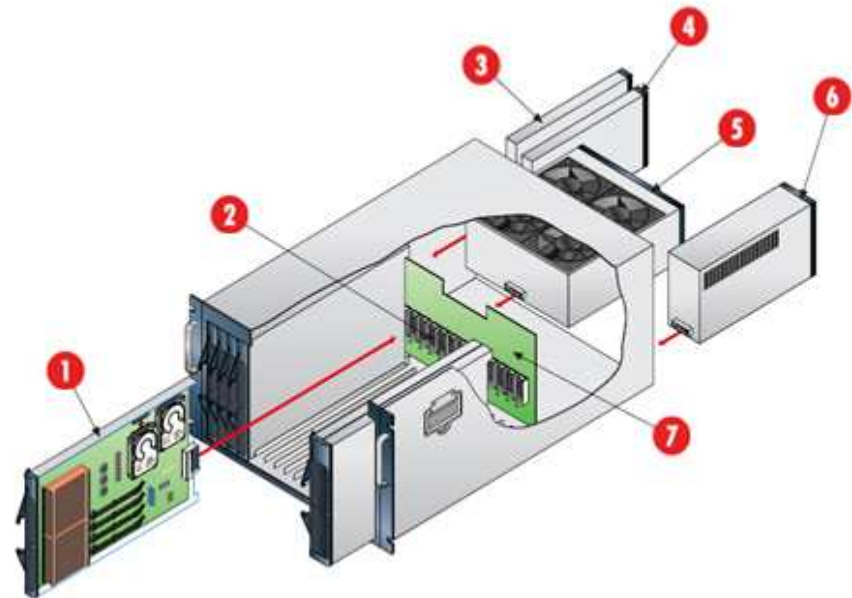
KR4 as the Backplane Channel Baseline

- Has already been proposed that the 40GBASE-KR4 channel be the baseline for 100G.
 - parthasarathy_01_0111.pdf
- Will lead to faster adoption of 100G in Blade Servers.

WHY?

Attributes of the Typical Blade Server

- **Modularity:** shared common resources (fans, power, management) simplifies the individual server module
- **Redundancy:** modules (blades, switches, power supplies) can be swapped live.
- **Midplane** is inherently part of the blade server chassis.
 - **Not a configuration option**
 - **Field upgrade is highly disruptive!**
- **Lifecycle:** Blade chassis typically lives for several generations of compute blade, supporting new server models 5-7 years after chassis intro.



The same midplane should work for all port speed configurations
...without undue cost burden when lower speeds used

Blade Server Channel Reuse:

What is the Legacy to Support?

A “Future Legacy” of 40GBASE-KR channels.

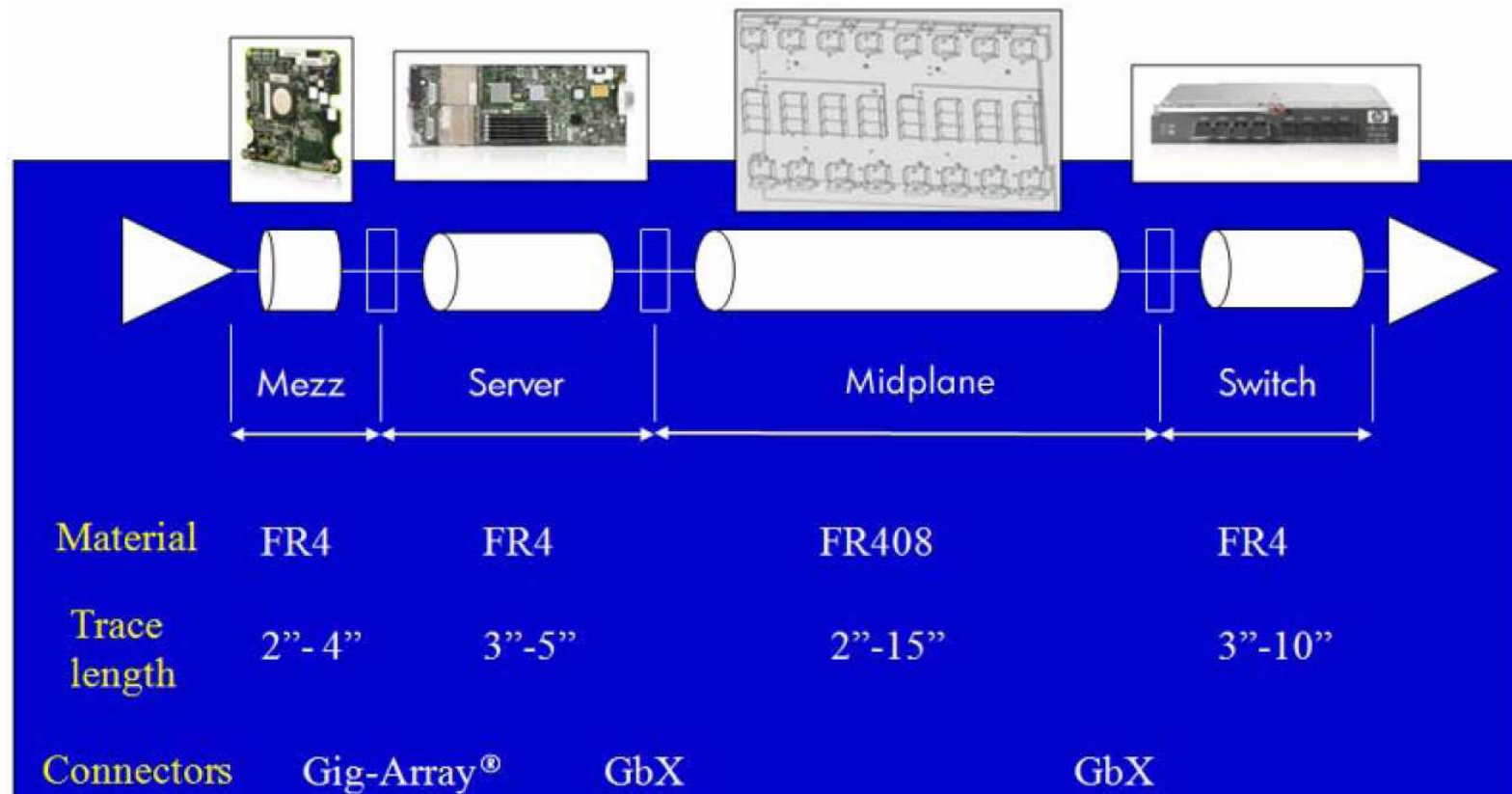
- Channels are already in the field, even though 40G products are not

Why a Four Lane Interface?

- Initial 10G deployment was XAUI & KX4 capable
 - >90% of blade server market by vendor unit share (IDC)
- 10GBASE-KR deployment came later...
 - Four lane interface maintained
 - Also used for other protocols like InfiniBand QDR
 - We can expect that systems not KR4 ready now will be upgraded soon

Four-lane, KR quality represents ~90% of blade server market by vendor unit share

Typical Blade Server Channel Topology



Source: HP.com

- Two basic mezzanine topologies:
 1. LAN signals route back to mother board, 3-conn model, ~60% by vendor unit share
 2. LAN signals direct to midplane, 2-connector model, ~30% by vendor unit share
- Same server & midplane support higher speed Ethernet by upgrading Mezz card & switch module

Balanced Cost Criteria – Blade Server Interconnect

25Gb NRZ quality channel

- Requires all new midplane designs
- Low Loss dielectrics 2-6x higher cost than basic FR-4
 - http://www.ethernetsummit.com/English/Collaterals/Proceedings/2011/20110222_WorkshopAM_Signaling_Goergen.pdf
- Channel for NRZ would burden every port, even though 100Gb attach is small.
 - Midplane and motherboard cost impact regardless of NIC speed used
 - Could force an entire motherboard to low loss material even if only a few inches are needed
 - 130-150sq.inches

KR4 quality channel (i.e., use multi-level signaling)

- Reuse existing infrastructure
- Modest enhancements could be incorporated in time, if required
- Maintain connector compatibility

Pay As You Go: add the cost of 100G only when 100G used, not on every server & midplane

Balanced Cost Criteria: PHY

Both Blade & Rack

- If the PHY is more complex, only burdens the 100G ports
 - ...but PAM4 PHY likely similar cost to NRZ PHY.
 - PHY is small part of The Big ASIC
 - KR PHY typically $\sim < 10\%$ of LAN Controller
- What if my Host Port should also support NRZ for an Optics interface?
 - Brute force approach would double the PHY
 - Actual implementations likely to find some reuse

In very approximate numbers:

Our choice is between adding 10% to the 100G Controller vs. 300%-600% to every blade server interconnect.

Development Environment of Volume Servers

- Most volume server designs outsourced to keep development costs low.
- CPU Platform Enablement distills complex design problems to design rules
 - Intel example: enabling for QPI, PCI Express Gen 3, SAS/SATA (6-8GT/s interfaces)
 - Enables many board designs with little simulation.
- That model unlikely to be successful with channel requirements for 25Gb NRZ.



Additional PCB Considerations for the Server Application

High Volume Server motherboards are very cost-sensitive.

- The PCB technology is mainstream, still predominately standard FR4.
 - Smartphone PCBs are much more hi-tech than servers!
- High speed, low loss PCB material is only 4.1% of the WW PCB production (by area) Prismark Printed Circuit Report, Q3'2010

Regulatory requirements impacting materials.

- RoHS: Restriction of Hazardous Substance
- Many compute products have made the transition to Lead Free
 - Though some exemptions are expiring
- Our next PCB challenge: **Halogen Free**
 - HF impacts the electrical properties
 - Best current HF material is similar to standard FR4... no Low Loss or Ultra Low Loss choices
 - International Electronics Manufacturing Initiative (iNEMI) pursuing this as a project
 - http://thor.inemi.org/webdownload/projects/ELSC/HFR-Free_PVC-Free_Timeline.pdf
 - Halogen Free was the only growing segment in 2009 with a +12% revenue growth vs -15% for the market at large.

Ultra Low Loss dielectric like Megtron-6 is not suitable for the future Server
Not Halogen Free and too expensive

The Broad Market Potential of Plug Compatible Ports

Plug Compatibility for Speed and/or Media is an attribute in common to many successful Ethernet port types.

Anything-BASE-T

- Four speed generations on a compatible jack
- New speeds often utilized initially at legacy speeds as market matures

Backplane Ethernet

- KX/KX4/KR all in use today
- Some product lines support all three in one blade system

SFP+

- Optical vs Direct Attach Copper very successful
- Most common 10Gb host port today
- Speed flexible 1G->10G



Consider backward (and forward) compatibility paths:

- 100G backplane should be plug-compatible with 40GBASE-KR4
- 100G external ports should be plug-compatible with 40GBASE-CR4 and short reach optical

→ A 100G NIC that also supports 40G & 10G is a very likely product scenario

Summary

- **Balanced Cost leads to:**
 - Put the complexity in the PHY
 - Maintain 40GBASE-KR4 channel
 - By extension... 40GBASE-CR4 channel
- **Maintain the Four-Lane interface**
 - Plug compatibility for backplane & external ports
- **Support Autonegotiation for multi-speed copper & backplane ports**