

BROCADE



40 Gigabit Ethernet Answers



Scott Kipp
Office of the CTO
skipp@brocade.com

With PHY support from:
Marco Ghisoni
Marco.ghisoni@zarlink.com
Zarlink Semiconductor

May 22, 2007

The April Meeting Had These 40GE Issues

What are the applications for 40G?

– Explain issues with LAG in the context of these applications

- Show more end user support
- What is the number and type of PMDs required?
- Must not delay 100G
- Don't want to double the R&D effort
- What is the relative cost between 4 x 10 G and 40G
- What is the basis for having two MAC rates vs one
- What is the relative time to market?

Copied from http://grouper.ieee.org/groups/802/3/hssg/public/apr07/frazier_02_0407.pdf

Bottom Line: 10GE has been so strictly defined that we need 40GE to bridge the gap between 1GE and 100GE



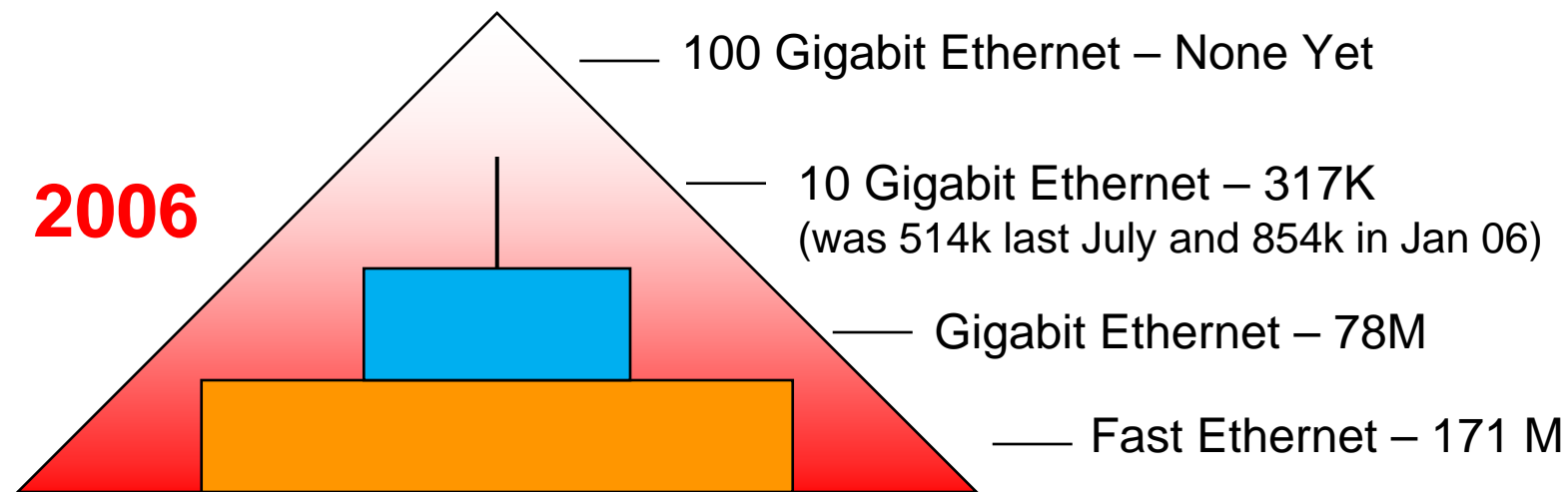
What are the applications for 40G?

High speed, cost effective data intensive applications.

The better question is who will buy it or use it?

Companies who need 10GE or more but not 100GE

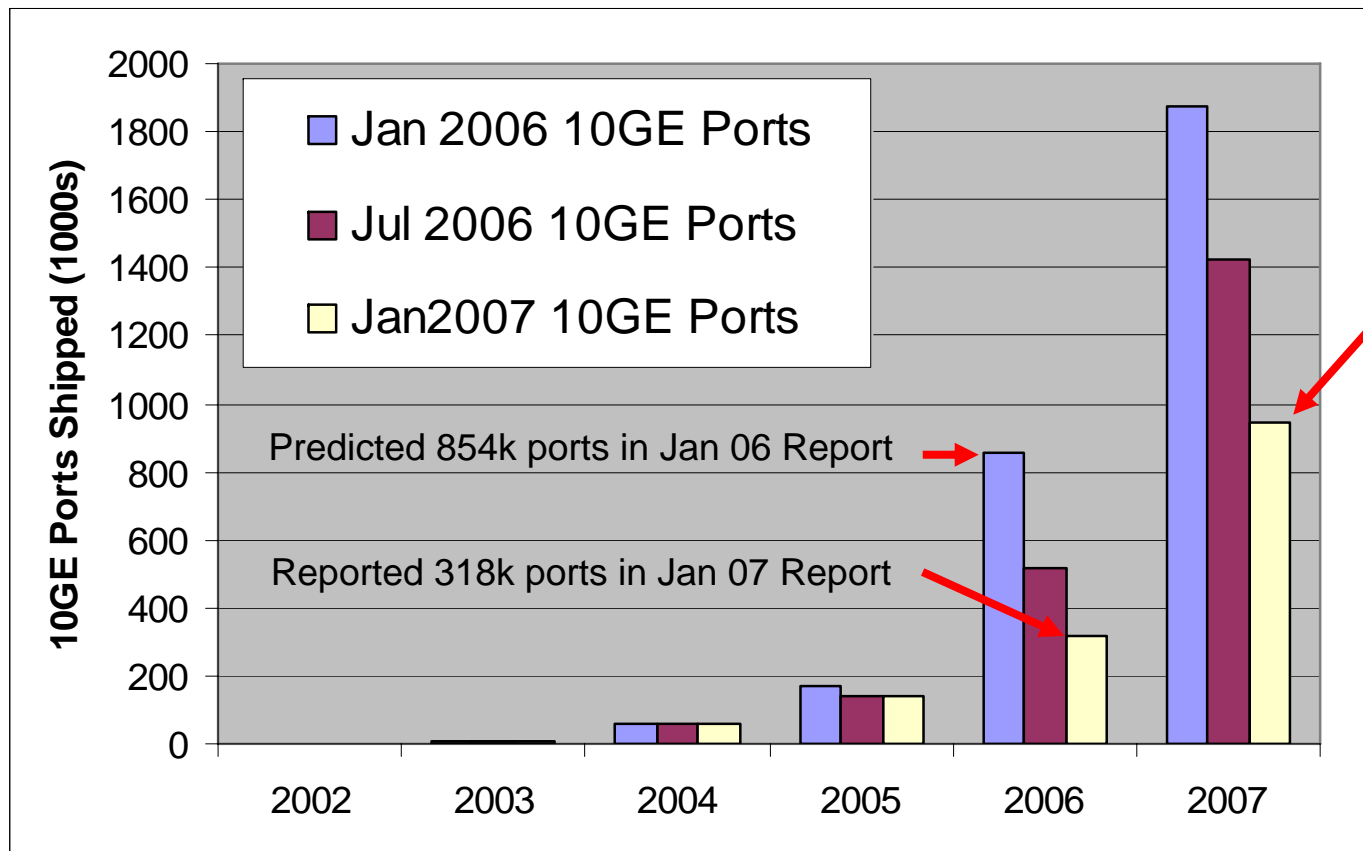
Much larger market than 100GE because it will be lower cost



Worldwide Switched Ethernet Port Shipments in 2006 according to Dell'Oro

Recap of 10GE Adoption

Dell'Oro reports and forecasts 10GE port shipments every six months. This graph shows forecasts and predictions from the three latest reports.



Forecasts for 2007 almost cut in half within last year

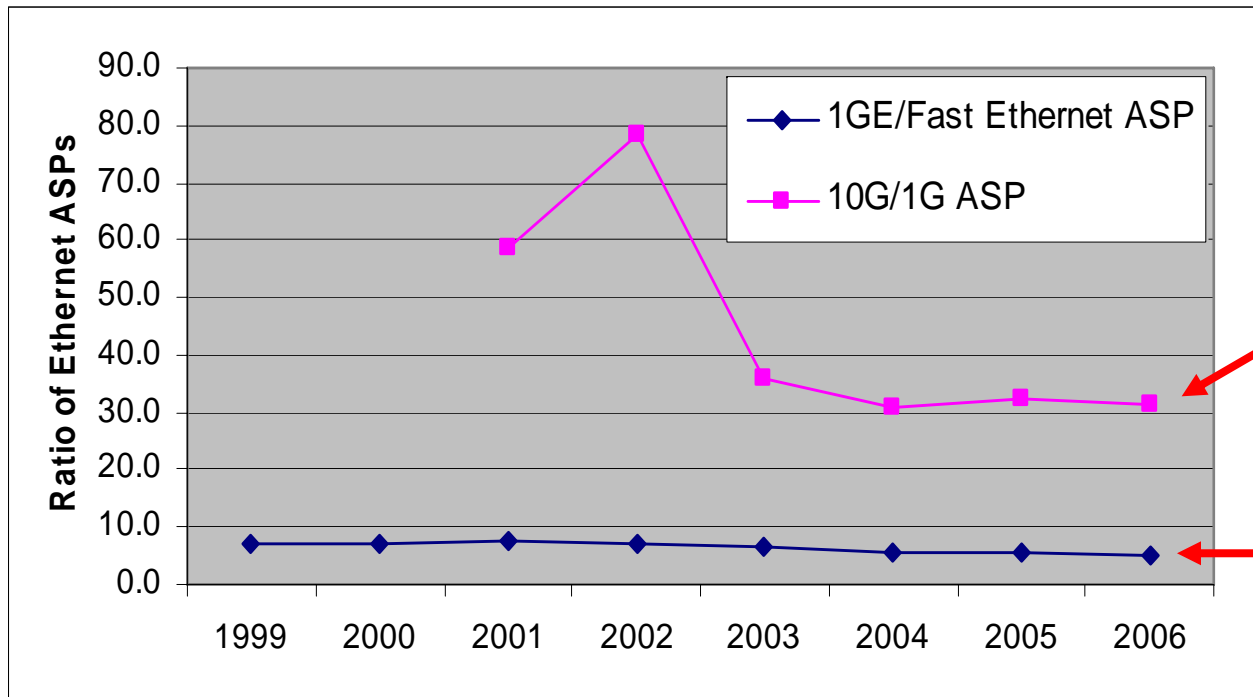
526k Ports shipped by Jan 2007
More ports should ship in 2007 than all previous years



10GE Can't Compete with 1GE

According to Dell'Oro, the average selling price (ASP) of a 10GE port was 31 times higher than a port of 1GE in 2006

- Price ratio between 1GE and Fast Ethernet was 4.8 times



10GE is in a different league than 1GE

1GE ASP approaching 3X Fast Ethernet and will largely replace it



Why is 10GE so slow (at being adopted)?

COST = **BIG CMF** ← Swiss Francs

Extract from Finisar proposal to 8GFC Fibre Channel
(<http://www.t11.org/t11/docreg.nsf/ufile/06-036v0>) :

Practical 10G VCSELs to date have been small aperture MM devices.

- Yield **VERY** poorly even to largest allowed spectral width
- This is largest single cost driver for 10GBASE-SR

Practical TX design almost impossible. Requires high ER and expensive yielding and testing

The market requirement to have 10GE support 300 meters and the rush to get 10GE out in 2001/2002 lead to problems in manufacturing and test.

- Some vendors have solved manufacturing problems
- All vendors have problems with testing and yielding to the specification
- 8GFC PHY is avoiding these pitfalls and is expected to be 4X cheaper than 10GE

Why is 10GBase-SR difficult to make?

10GBase-SR put great constraints on transceiver designs

- 300 meters * 10G = 3,000 MHz*km bandwidth length products (BWLP) of link
 - 1GE only has a bandwidth length product of ~1,000 MHz*km
 - OM3 fiber has a BWLP of 2,000 MHz*km, so the BWLP of the link exceeds that of the fiber by putting stress on the transceiver design
- 100 meters * 10G = 1,000 MHz*km BWLP of each lane for 100G or 40G
 - With a 66% reduction of the BWLP of the 10GBase-SR link, less stress is placed on the transceiver design.

The 3,000 MHz*km BWLP lead to many PMDs for 10GBase-SR

- 300 Pin MSA, XENPAK, X2, XPAK, XFP and soon QSFP
- Minimal volume production still and SFP+ is delayed until late this year or even next year

10GE Copper solutions are also stuck in the “difficult to make/use” category with a BWLP of 1,000 MHz* km since it spans 100 meters

- 10GBase-T – high power consumption at 6-15 Watts/port
- 10GBase-CX4 – cumbersome twinax pairs

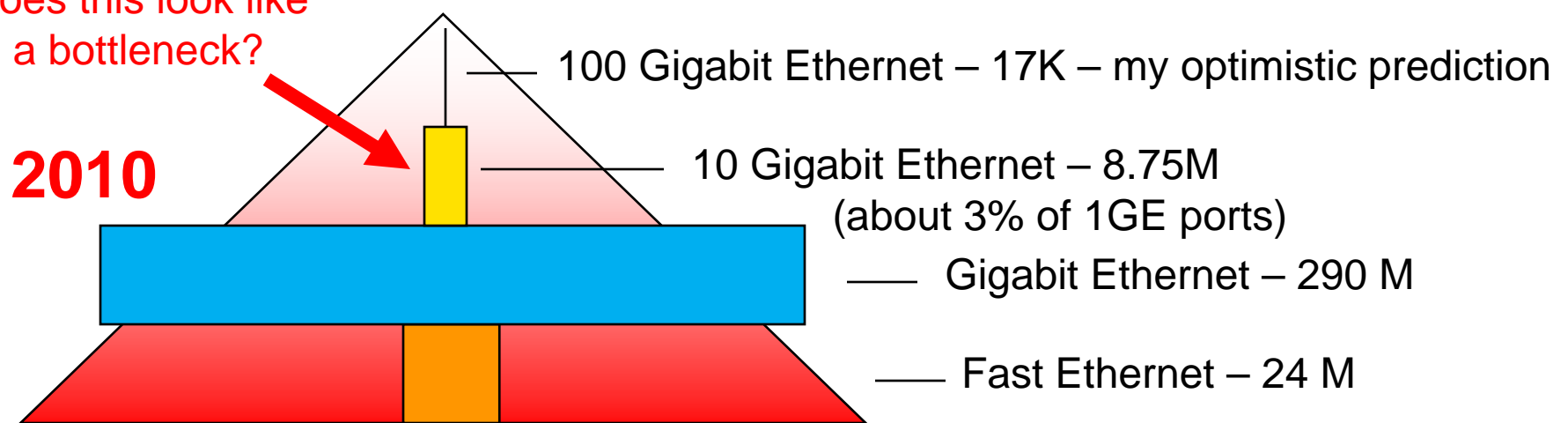


How many 100GE Ports will ship in 2010?

Dell'Oro hasn't predicted 100GE yet, but if 100GE follows a similar pattern as 10GE, then there will be minimal shipments in the first few years after the standard (This doesn't have to be!)

- One year after the 10G standard (2003), 10GE port shipments were 57,000 or ~0.2% of the GE port shipments
- 0.2% of 10GE Ports shipped would be 16,800 ports of 100GE in 2010 if the standard is released in 2009

Does this look like a bottleneck?



2010

Worldwide Switched Ethernet Port Shipments in 2010 according to Dell'Oro



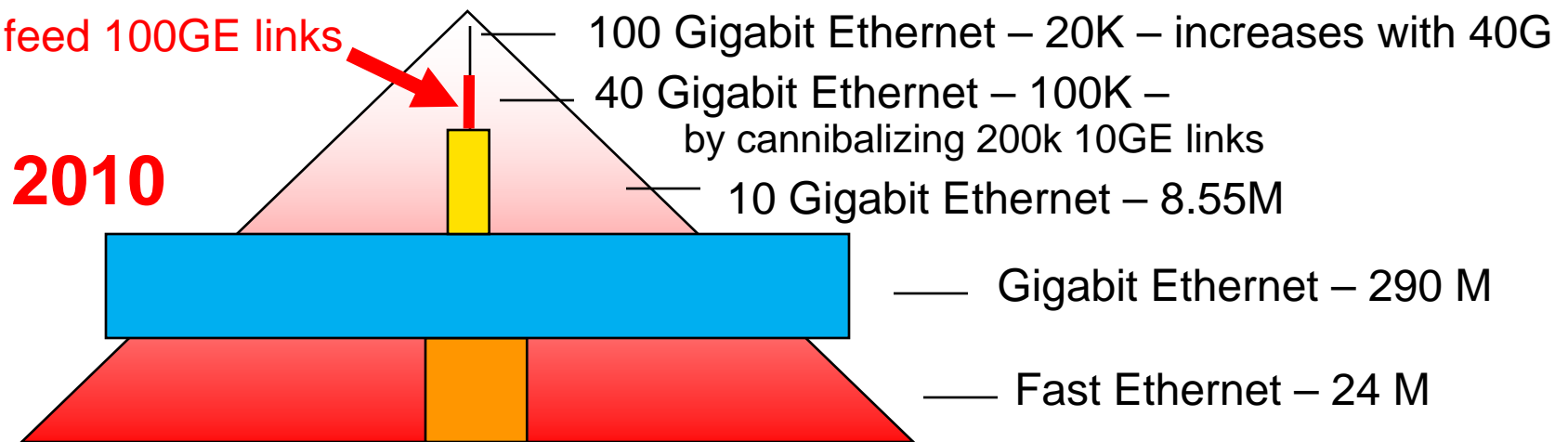
Where would 40GE fit in 2010?

40GE would compete with 10GE and could increase sales of 100GE

40GE could give better value than 10GE with only 33% more BWLP but 4 times the bandwidth

- Majority of new backbone links would be over new fiber in the core of the network, so PMDs for OM1 fiber (LRM like solutions) could be ignored.

40GE could relieve the 10GE bottleneck and feed 100GE links



10GE to Fast Ethernet Port Shipments in 2010 according to Dell'Oro



What is the relative cost between 4 x 10 G and 40G?

- Not the same if 40Gbase-SR4 is defined to go 100 meters while 10GBase-SR must traverse 300 m
 - Estimates are 10-50% less cost at 100 meters

Cost isn't everything. Users will buy 40GE over 4x10GE because of:

- Ease of Management – it's one port versus four
- Port Density – the QSFP is a factor of 3 denser than the 4 SFP+
- Cable Management – One ribbon is easier to manage than 4 duplex patchcords



Show more end user support

- United Airlines and IBM
- Brocade is an end user and we would like to use 40GE links between switches and to LANs to create File Area Networks (FAN)
- Most 40 GE customers are more cost conscience than 100GE users and can't afford to attend standards meetings
- 40 Gig is part of many other roadmaps
 - Fibre Channel Roadmap which also includes 10 Gig and 20 Gig
 - The Fibre Channel Roadmap can be seen at www.fibrechannel.org
 - OC-768 – telecom supports 40G today
 - InfiniBand roadmap supports 40 Gig - 4 lanes of QDR (10G/lane)
- All of these technologies support 2x, 3x or 4x jumps in the roadmap,
 - Why does IEEE insist on these 10X leaps of faith?

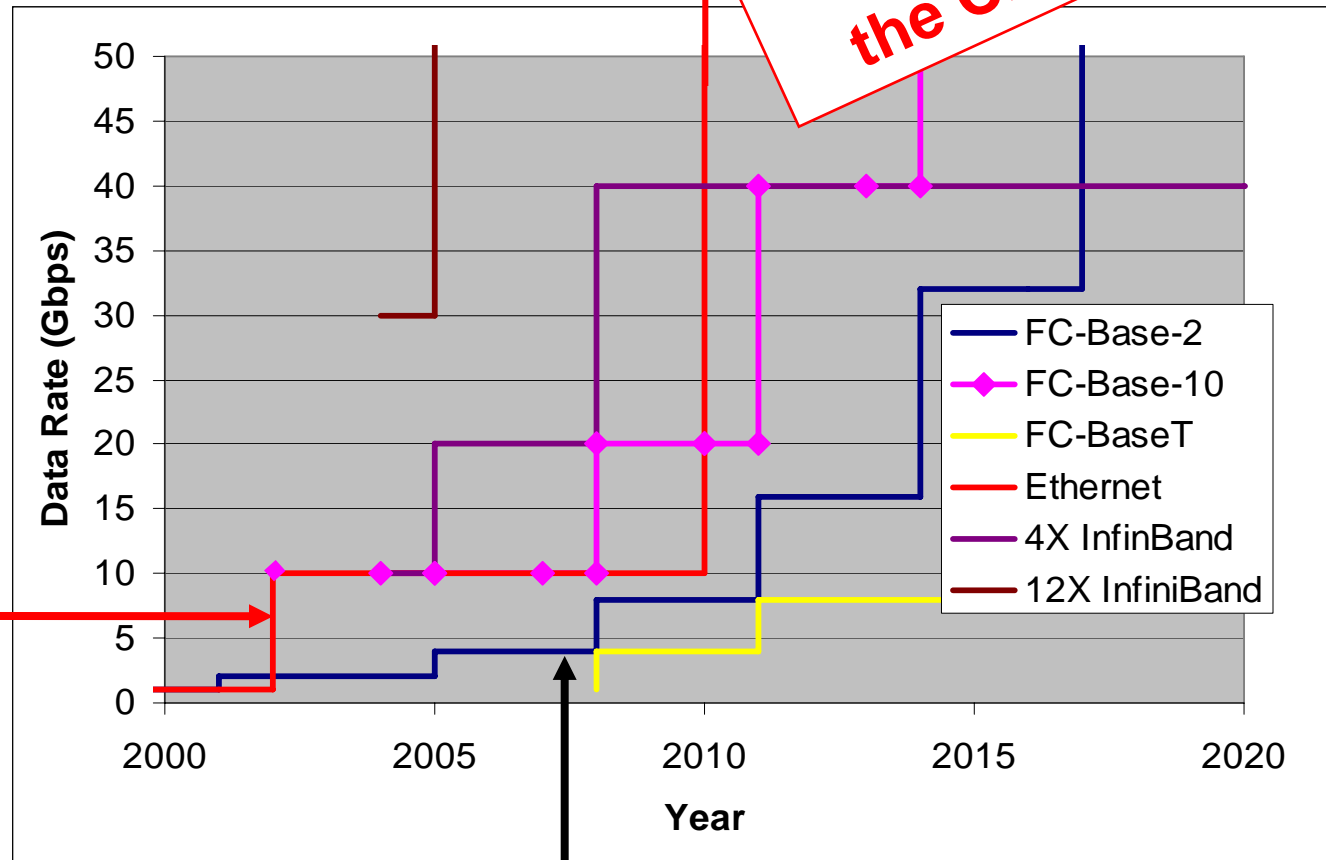


Comparison of Gbps Roadmaps

The second leap of faith of 90G to 100GE will not be easy.

100 GE Off the Chart!

The last 10X leap of faith to 10G has fallen short up until now. Fibre Channel jumped too!



Fibre Channel's tortoise like 2X progression has 4GFC outshipping 10GE by 10x at 1/10th the ASP!

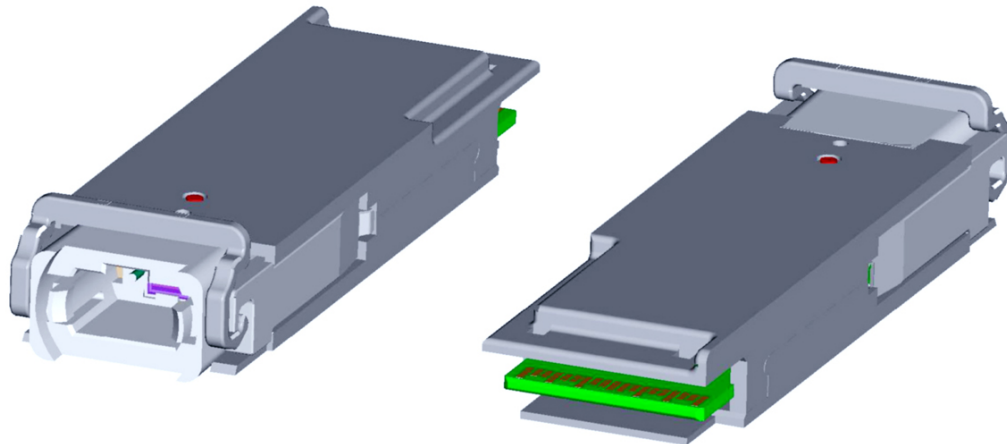


What are the number and type of PMDs required?

- Same as 100GE
 - Must skip the 10GE PMDs designed for OM1 Fiber (LRM)

Another question is what transceivers are being standardized?

- For multimode applications, the QSFP should work fine, is complete and is in pre-production at 20Gig (4 X 5G / lane)
 - **A 40G QSFP (4 X 10G / lane) availability announcement is imminent this year**



The QSFP:
Find out more at
www.qsfpmsa.org

- For singlemode applications and multimode applications, the X40 should work well and is being standardized right now
- What Transceivers are being standardized for 100GE? 0

Must not delay 100G

If we continue talking about 40GE, it will delay 100 GE

Agreeing to add 40GE instead of endlessly debating will lead to faster standards completion.

The better question is will 40GE delay the standard significantly or even measurably.

Sun showed that there is minimal effort to standardize 40GE and should not measurably delay 100GE:

See http://www.ieee802.org/3/hssg/public/apr07/muller_01_0407.pdf

We're already looking at standardizing 10 lanes at 10 Gig and 4 lanes at 25 Gig. 40GE would require 4 lanes at 10 Gig with minimal effort



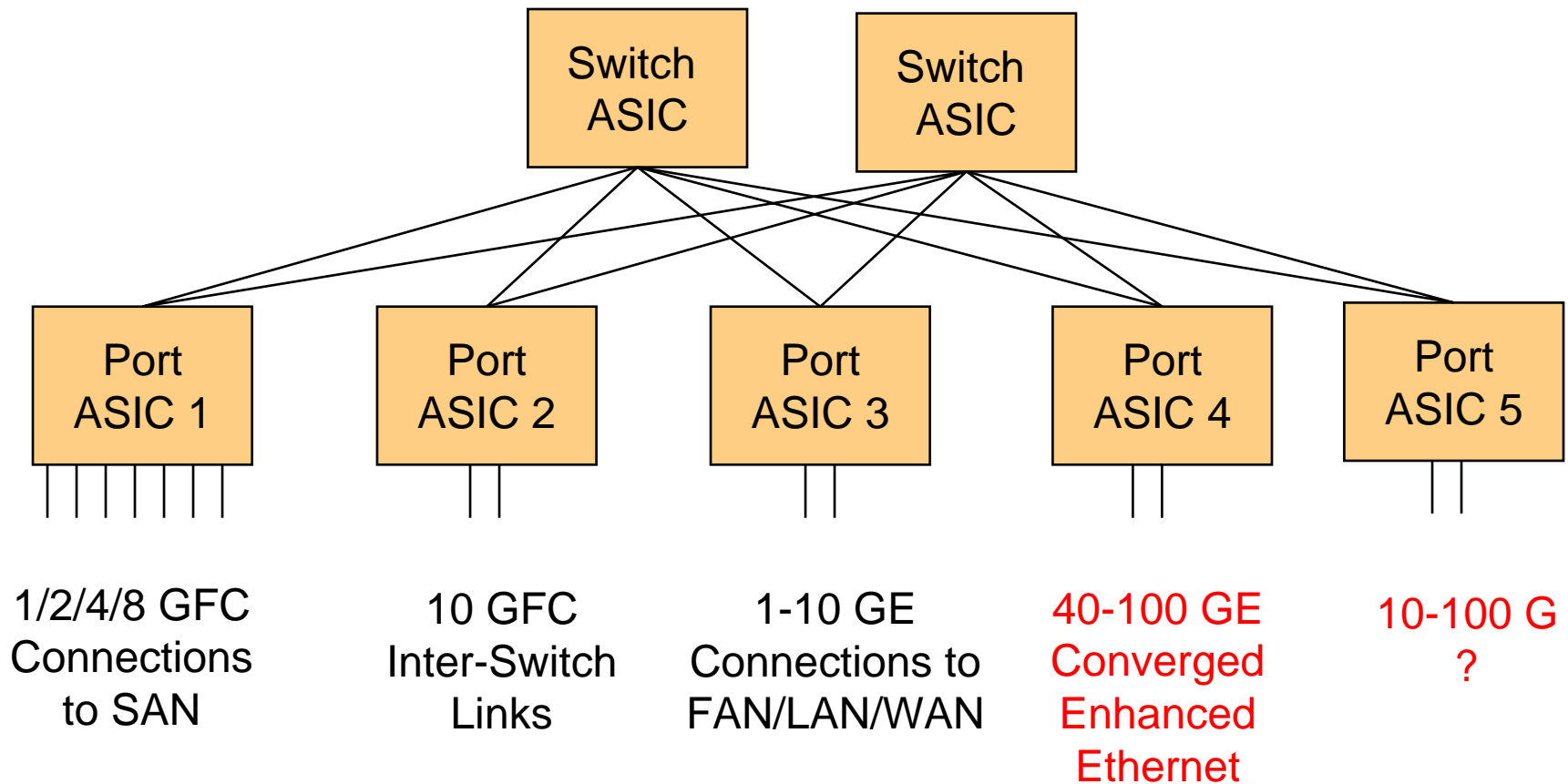
Don't want to double the R&D effort

- Optimizing R&D effort is definitely a good thing
 - Did we learn anything from 10GE?
 - Should we have a transceiver MSA before the standard is complete?
 - Are the components cost-effective and manufacturable?
 - 40GE doesn't need to double R&D Effort but new chips will need to be developed
- Development at the 40G optical layer is almost complete for some companies
 - Some companies are still researching it
- 40GE will be an intermediate step towards 100GE that will aid in quicker development of 100GE transceivers



What is the basis for having two MAC rates vs one?

- Because it's being done or **will be done** at Brocade...



Mind the MAC GAP (Between 1GE and 100GE)

- We need a 40G MAC/Ethernet to reach a larger market that is underserved by expensive 10GE products and overserved by 100GE products.
 - Factors of 10 jumps do not fit as many applications as incremental changes
 - One Size Fits All - fits very few well
- Interoperability concerns for matching 40GE and 100GE are unfounded. Ethernet is frame based and resolving lane mismatch has been accomplished by many vendors.
- For mission critical applications, links usually have to be redundant
 - Companies that need redundant 100G links will have to pay at least twice as much as companies that want redundant 40G links



What is the relative time to market?

- Sooner than 100GE because of the physical layer
 - 40 Gig OC-192 PHY available today
 - 40 Gig QSFPs coming this year for multimode applications
 - 40 Gig QSFP short reach (2km) singlemode possibly in 2008
 - 4X10 Gig X40s for multimode and singlemode applications should be available next year
 - These are well before any standard 100GE transceivers



Why 40GE?

- Because it is here at the physical layer and can be here sooner than 100G
- Because it fits the needs of the customer better than a 10X jump
 - Most cost conscience customers can't afford to come to these meetings and ask for 40GE!
- Because it will have 1.8-2.2x less cost than 100G
 - 10GE has left a gap between 1GE and 100GE
 - The best thing to increase demand for 100GE is to redefine 10GE in a cost conscious manner that spans less than 100 meters
- Because it will have minimal impact on the 100GE standard and it could improve the delivery time of cost effective solutions
- Because it matches up to supporting technologies that grow at 2X and 4X rates



Two Visions of the Future

1. If 10GE is not redefined and remains uneconomical at 300 meters, then 40GE solutions will be considered high value and can cannibalize the 10GE market. 40GE will fill the bandwidth gap in large data centers and drive bandwidth to the 100GE core.
2. If 10GE is redefined to make it economical to less than 100 meters, 10GE could effectively drive demand to 100GE and 40GE would not be needed.

With neither of these approaches, there is a gap between 1GE and 100GE.

