Higher Speed Ethernet: 100 *or* 40 GbE?

Drew Perkins
dperkins@infinera.com
408-572-5308

Legal Disclaimer

Although the forward-looking statements included in the presentation may constitute “forward-looking statements” within the meaning of the Private Securities Litigation Reform Act of 1995, they are in fact merely the musings of a madman.

The Parties acknowledge and agree that the information set forth herein (the “Future Information”) is provided to the recipient for informational purposes only and that any dates accompanying Future Information is provided as target dates only. The inclusion of such Future Information and dates is in no event to be interpreted or construed as an obligation on the part of Infinera or the Networking Industry to provide such Future Information, either in whole or in part, or in any particular manner or time frame.

There shall be no remedy or recourse by recipient against Infinera or the Networking Industry if either fails to deliver such Future Information in accordance with the target dates or at all. The Parties further acknowledge and agree that to the extent Future Information is developed in the future by Infinera or the Networking Industry and made available to recipient, any purchase thereof by recipient would be subject to separate negotiation as to any terms and conditions. No such terms and conditions have been agreed to date. Further, the recipient is not obligated to test or purchase any Future Information.

Service Provider Requirements For Higher Speed Ethernet

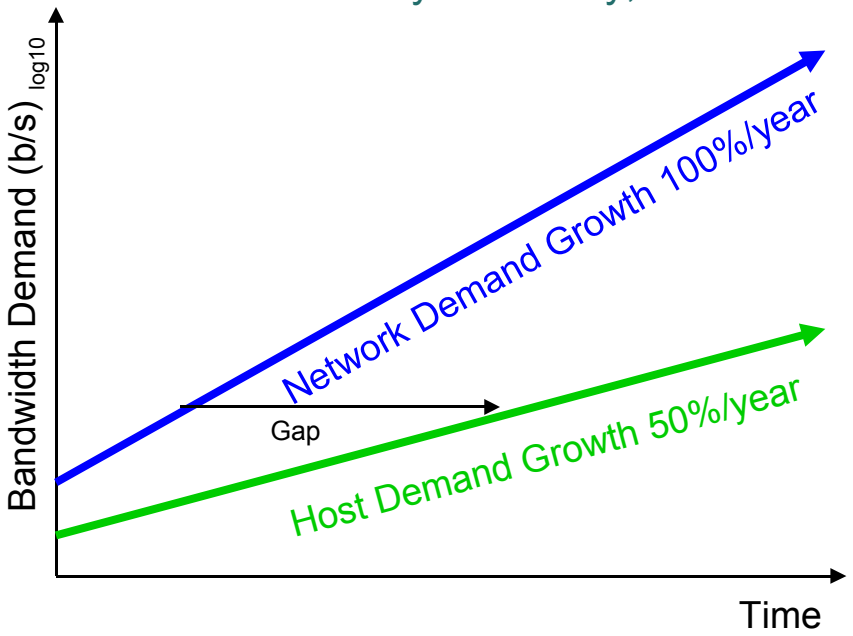
- Presentations from individuals associated with AMS-IX, Comcast, Cox, DT, EDS, Equinix, Google, KPN, LBNL, NTT America, NYSE, Sprint, TWC, Yahoo!, etc.
- Experiencing ~2x growth/year
- Internet core links use Nx10G (N ≤ 8 now, 16 soon) today
  - 80-160 Gb/s
- 100 GbE imperative NOW!
  - Tb/s links required by 2010 (before 100 GbE available?)?
- (DWDM) Transport network interoperabilty
- Carrier class OAM capabilities
  - Fault isolation of every link: AIS, etc.
- Hitless, incremental growth and graceful degradation

Server and End-Station Host Requirements

- Bandwidth growth rate commensurate with processor speed and interconnect (e.g. PCI-Express, Infiniband) performance growth rates
  - Moore's Law: 2x increase every 18-24 months
  - PCI-Express Gen 2 emerging now
    - Gen 2 (DDR) lanes run at 5 Gb/s encoded (4 Gb/s unencoded)
  - 8x PCIe2 ≤ 32 Gb/s, so 40 GbE pipes can't be filled
  - 16x PCIe2 ≤ 64 Gb/s, so why do hosts need 100 GbE?
- 40 GbE right-sized 2012-2015
- Ultra low cost and high volumes
  - Many intend for Ethernet to compete with and displace Infiniband and Fiber Channel
    - E.g. Converged Enhanced Ethernet (CEE)
- Blade servers upgradable from 10 GbE (4x 3.125 Gb/s) to 40 GbE (4x 10.3125 Gb/s) without chassis backplane upgrade

The Higher Speed Ethernet Dilemna

- Network interconnections (switch-to-switch)
  - 10G (OC-192 POS) and Nx 10 GbE has been standard for years
  - 40G (OC-768 POS), Nx40G beginning deployment now
  - **100 GbE** is imperative for the network NOW!!!
- Host connections (server/PC-to-switch)
  - Servers beginning mass adoption of 10 GbE only now
  - **40 GbE** won't be necessary for a few more years, while 100 GbE unnecessary until 2015?
  - Laptop PCs with 10 GbE still years away, while 40 GbE is possibly a decade away



Higher Speed Ethernet: 40G or 100G?

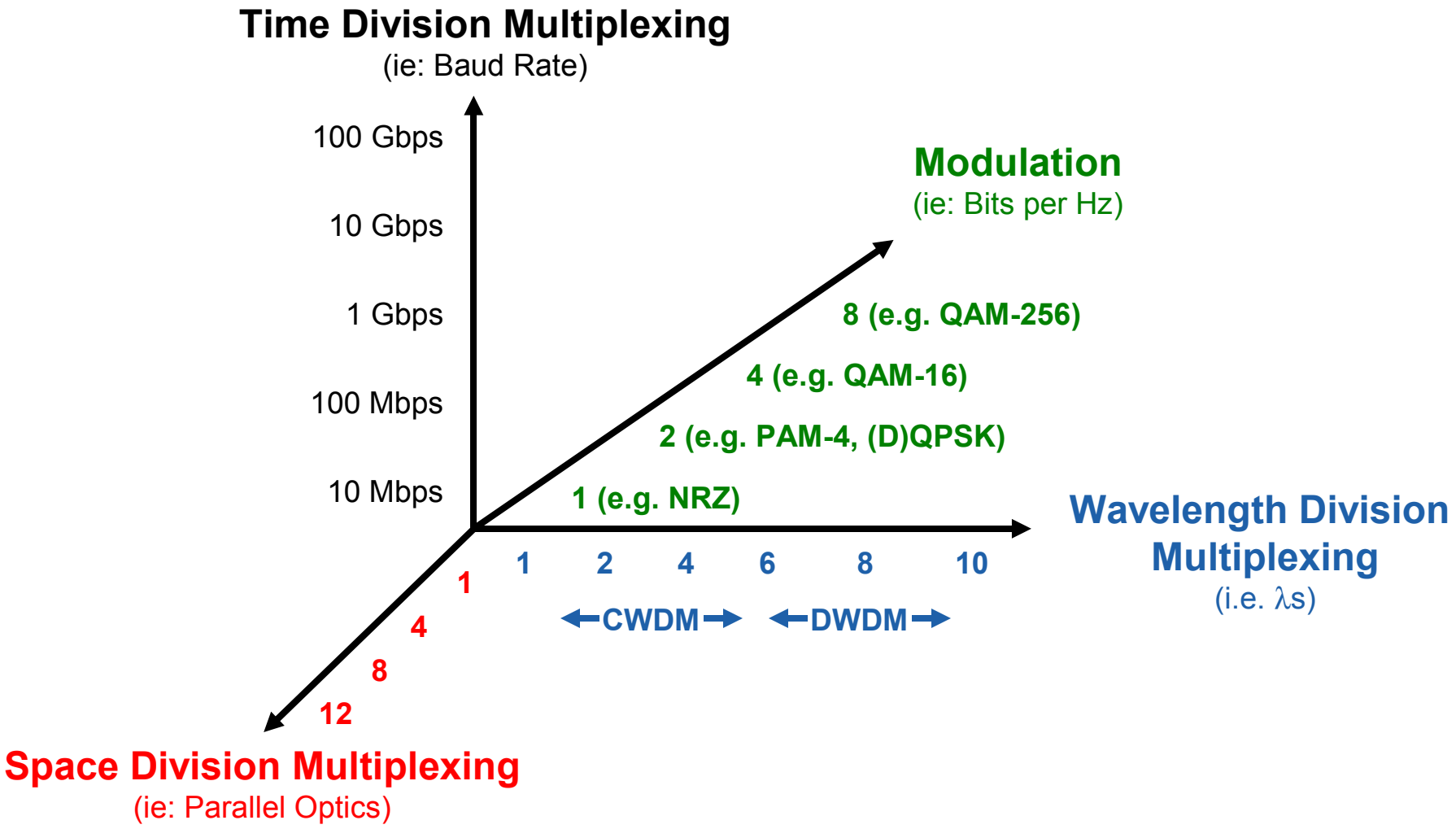
- After months of debate, the answer *may* be:

100 GbE for the Network, AND
40 GbE for Hosts!

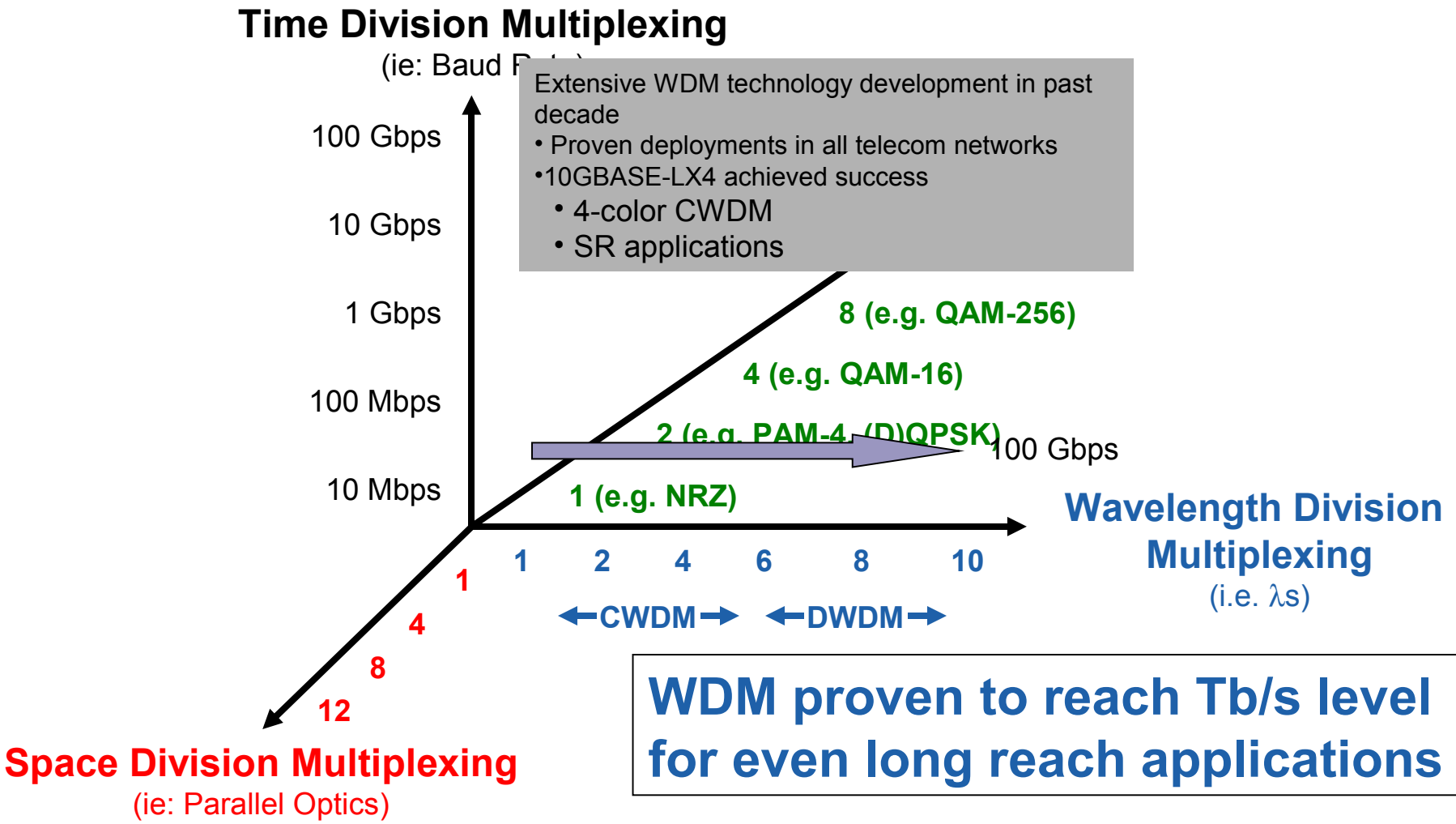
100 GbE PMD Objectives (Approved)

- 10 m on a copper cable assembly
- 100 m on OM3 Multi-Mode Fiber (MMF)
- 10 km on Single-Mode Fiber (SMF)
- 40 km on SMF

Technological Approaches to 100 GbE



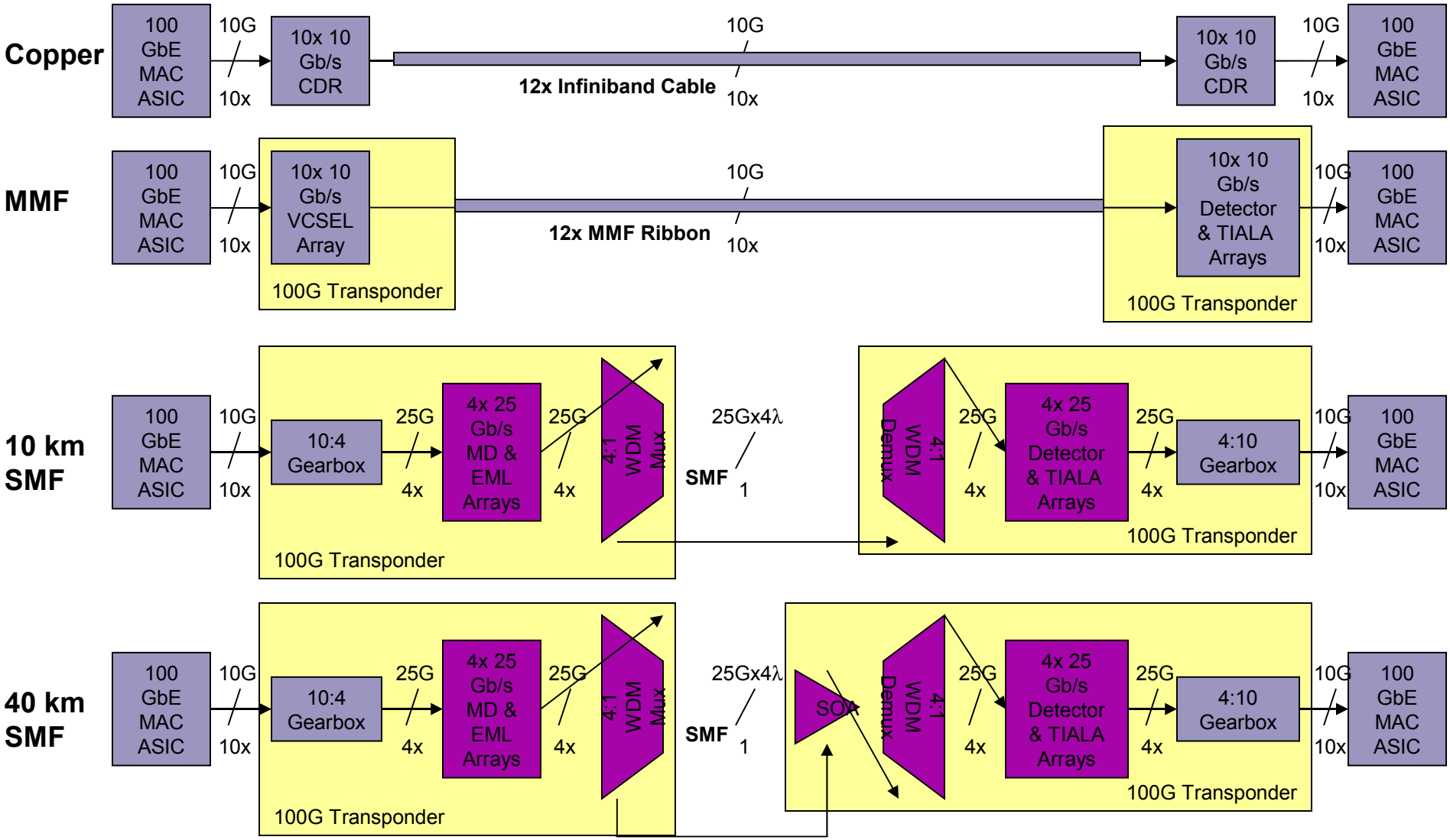
WDM Proven Beyond 100 Gb/s



Proven 100 GbE PMD Technical Feasibility

- 10 m on a copper cable assembly
  - 10x 10 Gb/s over Infiniband/CX4-style coax
  - 10GBASE-CX4 runs at 4x 3.125 Gb/s
- 100 m on OM3 Multi-Mode Fiber (MMF)
  - 10x 10Gb/s using 850 nm VCSEL arrays and fiber ribbons
  - 10GBASE-S runs at 1x 10 Gb/s using 850 nm VCSELs
- 10 km on Single-Mode Fiber (SMF)
  - 4x 25G using 1310 nm EMLs (possibly DMLs) and WDM over single fiber pair
  - 10GBASE-LX4 runs at 4x 3.125 Gb/s using 1310 nm DMLs over single fiber pair
  - 10GBASE-L runs at 1x 10 Gb/s using 1310 nm DMLs
- 40 km on SMF
  - 4x 25G using 1310 nm EMLs, Semiconductor Optical Amplifiers (SOAs) and WDM over single fiber pair
  - 10GBASE-E runs at 1x 10 Gb/s using 1550 nm EMLs
  - First use of SOA technology in any standard (risk?)
- Notes:
  - No 100 Gb/s serial PMDs
  - No metro/LH DWDM interfaces

Likely 100 GbE PMD Architectures



Proposed 40 GbE PMDs

- 1 m on backplanes
- 10 m on a copper cable assembly
- 100 m on OM3 Multi-Mode Fiber (MMF)

Proven 40 GbE PMD Technical Feasibility

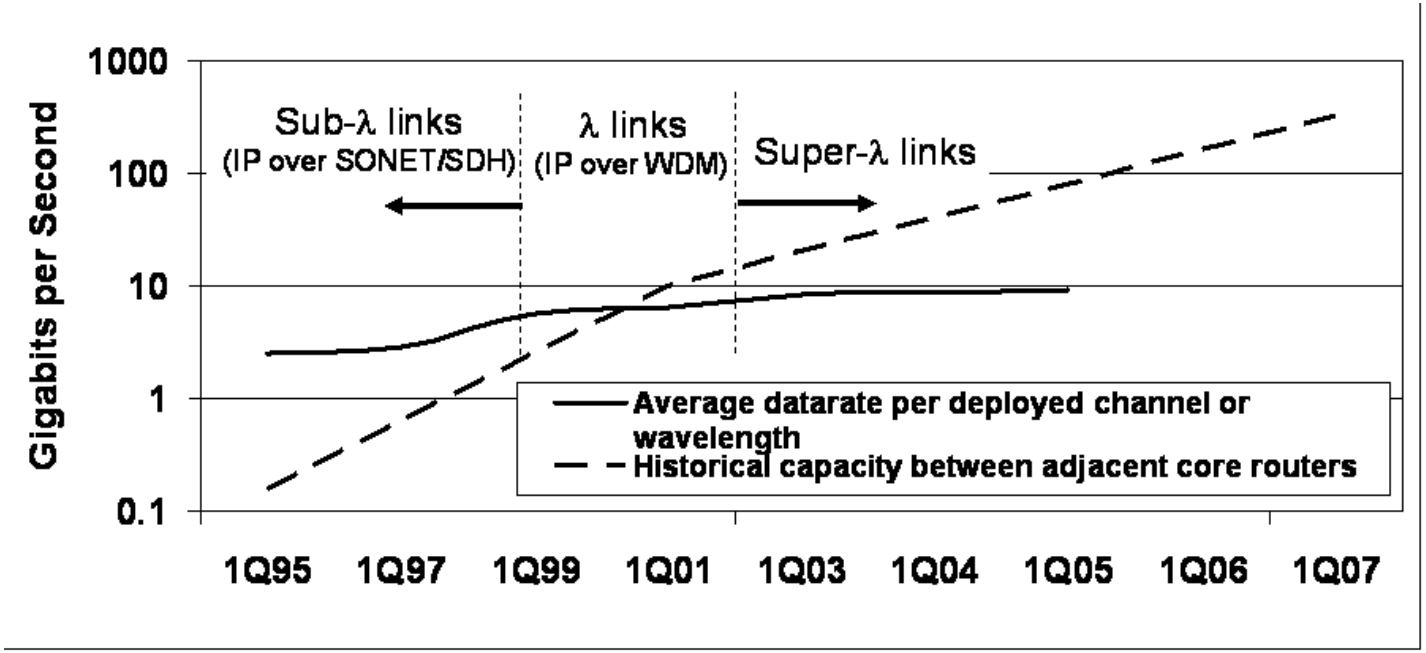
- 10 m on a copper cable assembly
  - 4x 10 Gb/s (instead of 10x 10 Gb/s for 100 GbE)
- 100 m on OM3 Multi-Mode Fiber (MMF)
  - 4x 10 Gb/s (instead of 10x 10 Gb/s for 100 GbE)
- Notes:
  - No 40 GbE SMF PMDs?
  - No 40 Gb/s serial PMDs?
  - No metro/LH DWDM interfaces

How Will 100 GbE Be Transported Over DWDM?

- 100 Gb/s serial?
  - ITU-T working on adding new ODU4 rate to G.709
    - Proposals for 100 Gb/s payloads
  - Dual polarization DQPSK being studied
    - 4x 26.75 Gb/s channels (with 7% FEC)
- 3x 40 Gb/s?
  - ITU-T G.709 already specifies ODU3-3v
    - 3x 40 Gb/s
  - ODU3e being considered
    - 3x 41.25 Gb/s (with 64B/66B)
  - 20% wasted bandwidth
- 10x 10 Gb/s?
  - ITU-T recently relented and specified OTU2e/ODU2e
    - Overclocked OTU2/ODU2
    - 10x 11.1 Gb/s (with 7% FEC)

Drivers for a Super-λ (Multi-wavelength) Protocol

- Growth of IP links historically and dramatically out-paced capacity of a single wavelengths



- Bandwidth requirements strongly favors approach leveraging multiple wavelengths (aka Super-λ service, composite links, LAGs, etc.)

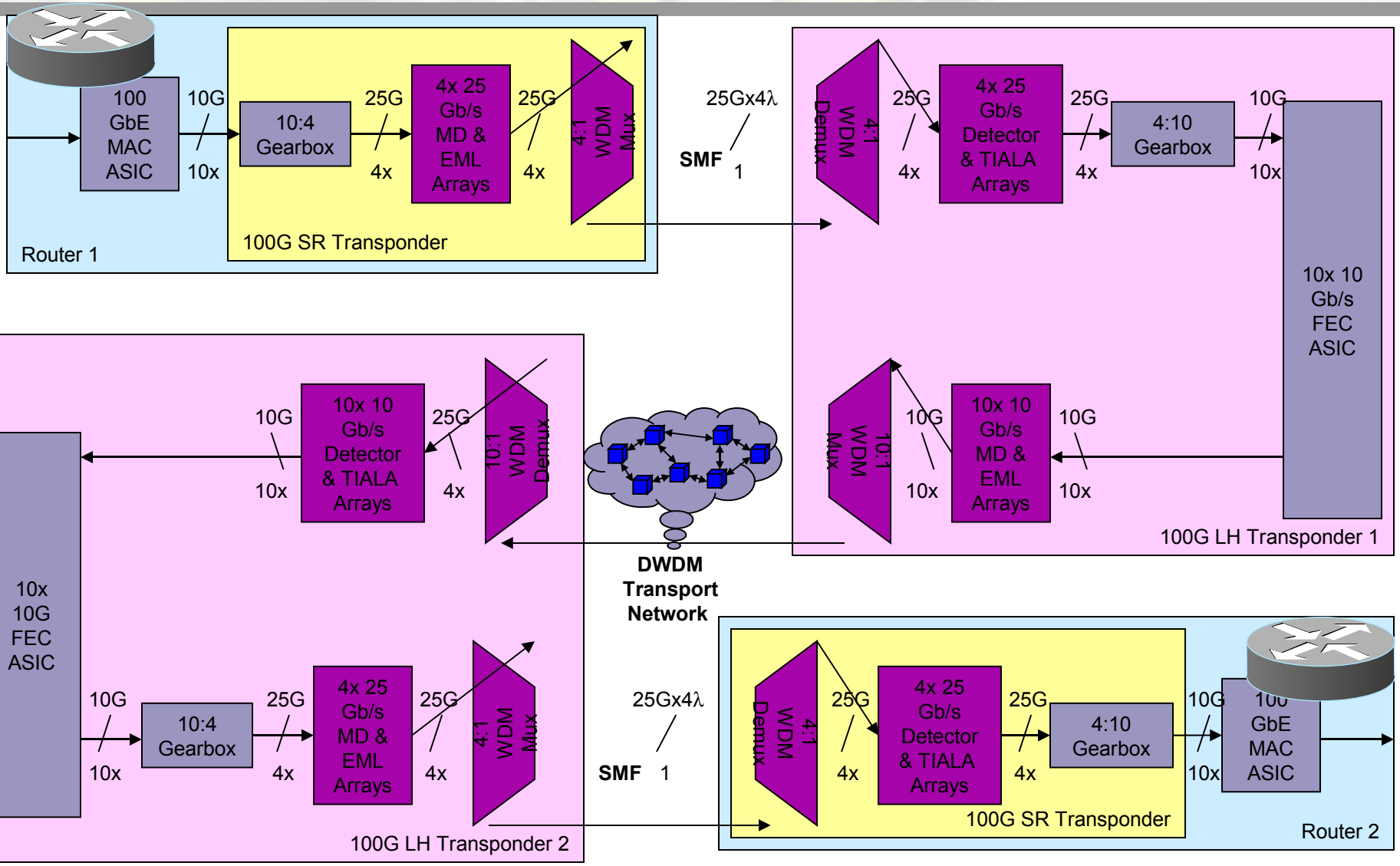
It's Really a Question of Economics

- OC-768 service provider experience
  - 4x bandwidth increase, but » 4x cost increase
  - Not 2.5x as historically experienced and now expected/wanted
  - I.e. 1x 40 Gb/s » 4x 10 Gb/s
    - 3x 40 Gb/s » 12x 10 Gb/s > 10x 10 Gb/s
- Key question: Will 100 GbE cost « 10x 10 GbE?
  - Some predict 100 GbE may cost ~2x 40G POS
- 10x 10 Gb/s λs may remain « 1x 100 Gb/s λs for some time

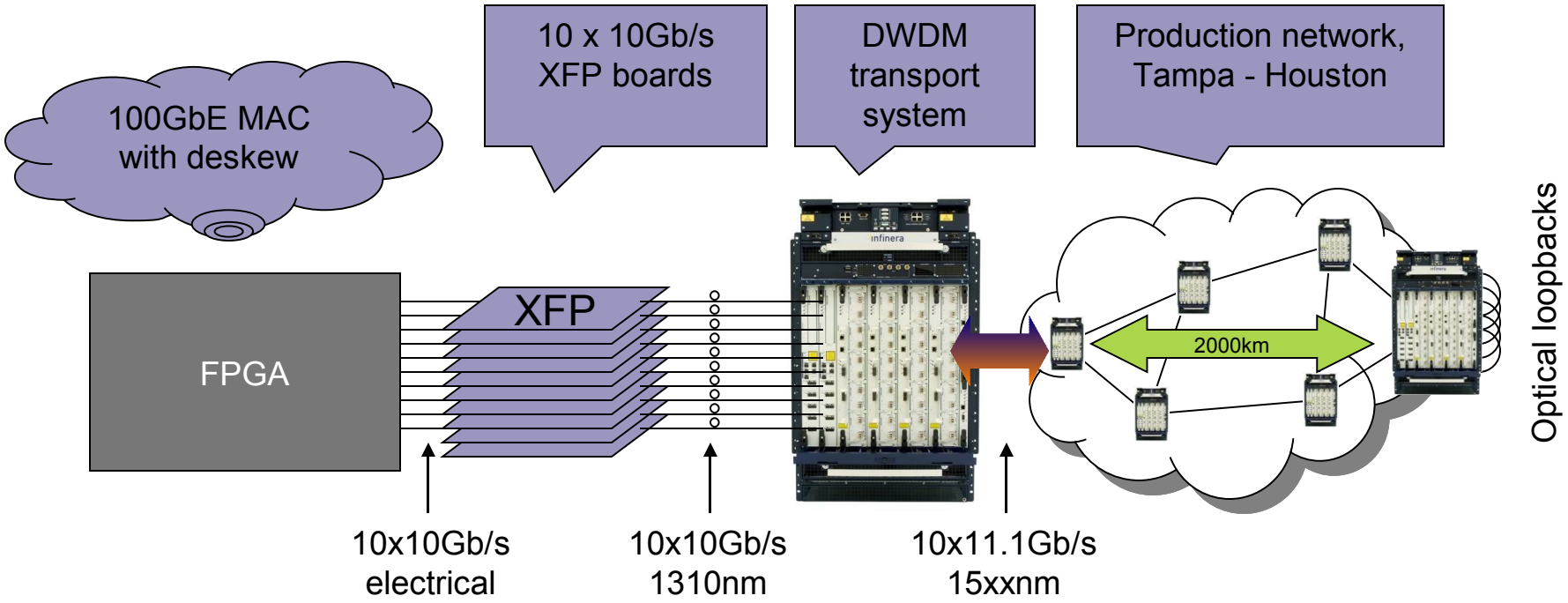
Possible Channel Bonding Techniques

- Traffic may be distributed over multiple links by a variety of techniques
  - Flow Distribution
    - I.e. LAG/ECMP
    - IP source/destination based distribution
    - MPLS labels problematic
  - Bit/Octet/Word Distribution
    - In a manner similar to 10GBASE-X, increments of the serial stream are assigned sequentially to lanes
    - Minimal additional overhead required to allow re-alignment at the receiver
  - Packet Distribution
    - In a manner similar to ML-PPP, sequence numbers added to packets to enable re-ordering at the receiver
    - Large packets within the stream may induce excessive delay/delay variation to smaller, latency-sensitive packets
  - Packet Distribution with Fragmentation
    - Fragmentation bounds buffering requirements and delay associated with packet size and packet size variation
    - Overhead/link inefficiency is a function of the maximum fragment size chosen
    - At 100 Gb/s and above, a fragment size can be chosen such that an effective compromise between link efficiency and the QoS of individual, time-sensitive flows can be readily achieved
- HSSG has heard presentations on all of the above, but approach won't be chosen until Task Force formed

100 GbE Over DWDM Transport –
10x 10Gb/s Example (1 of Several Possibilities)



Recent 100 GbE Over DWDM Transport Demos



- 100 GbE over 10x 10G DWDM demoed at SC2006, November, 2006
- 100+ Gb/s serial experiments and demonstrations announced at OFC 2007 and ECOC 2006

Thanks!

Drew Perkins
dperkins@infinera.com
408-572-5308