Ethernet Technology Summit

What is FlexEthernet?

Mark Gustlin & Faisal Dada – Xilinx

Ethernet Technology Summit

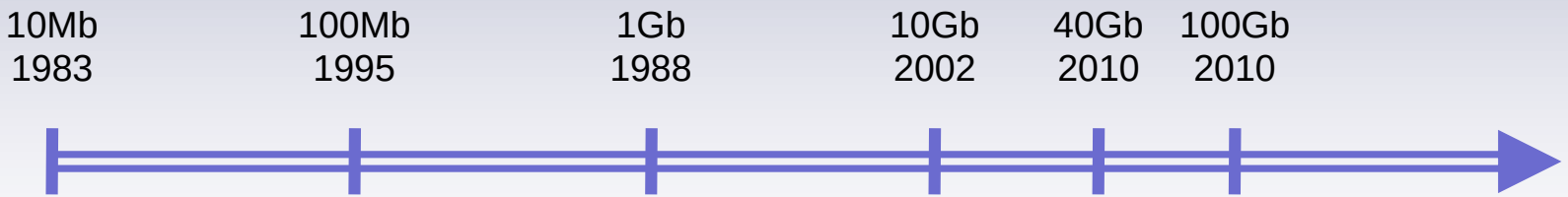
Outline

- Background on Ethernet Speeds
- FlexEthernet use cases
- Overview of FlexEthernet

Ethernet Technology Summit

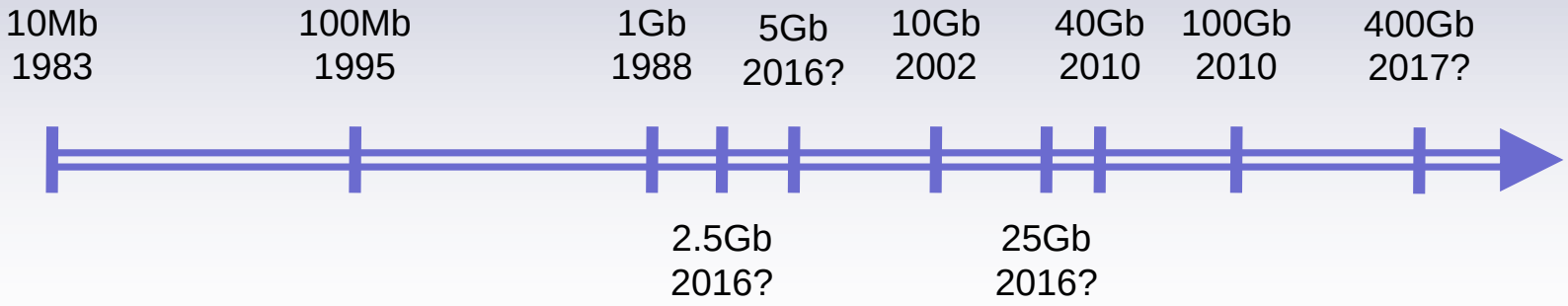
“Traditional” Ethernet

- Traditionally Ethernet went in steps of 10x, starting at 10Mb
- That long held trend had a wrinkle with 100/40GbE in 2010
- 40GbE was the first market optimized speed, for servers
- 40GbE was the first time we realized that cost drivers would demand a speed optimization for a market



Today's Ethernet

- There are many speeds being defined, even in between established rates
- New markets are demanding cost and power optimized solutions
- 2.5/5GbE are optimized for WAP and existing cable infrastructure
- 25GbE is for TOR to servers, optimizing cabling infrastructure
- There is also a consortium defined 50G speed



Ethernet Technology Summit

Tomorrow's Ethernet

- Expect more speeds as new markets require optimized solutions
- Faster Ethernet becomes more and more difficult as technology is pushed
  - Don't expect to see 10x jumps in speeds again
- FlexEthernet is one new area of Ethernet innovation
  - Rest of the presentation explores that area

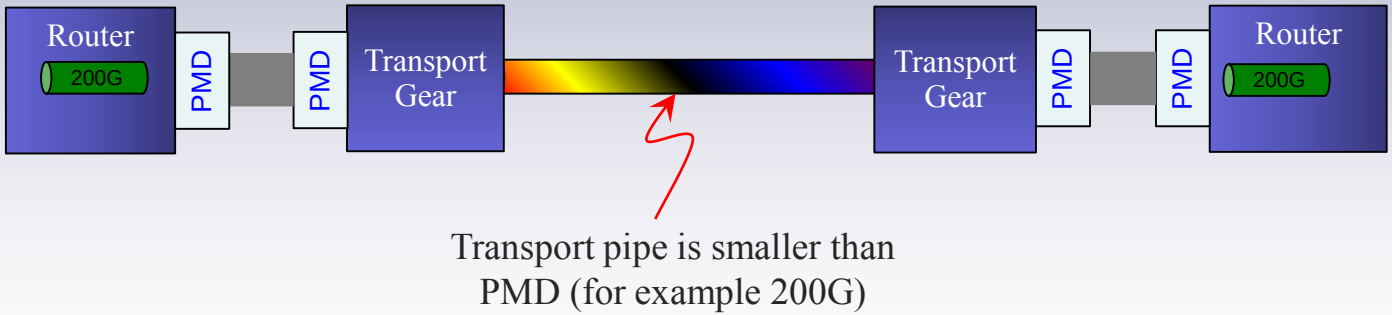
Ethernet Technology Summit

Why FlexEthernet?

- Router and transport gear is now evolving at different rates
  - Expensive transport gear may not be optimally used with fixed rate interfaces
- Desire to support simple transport of n x Ethernet streams across a faster interface
- Provide a more efficient ‘LAG’ mechanism

Router to Transport Interface

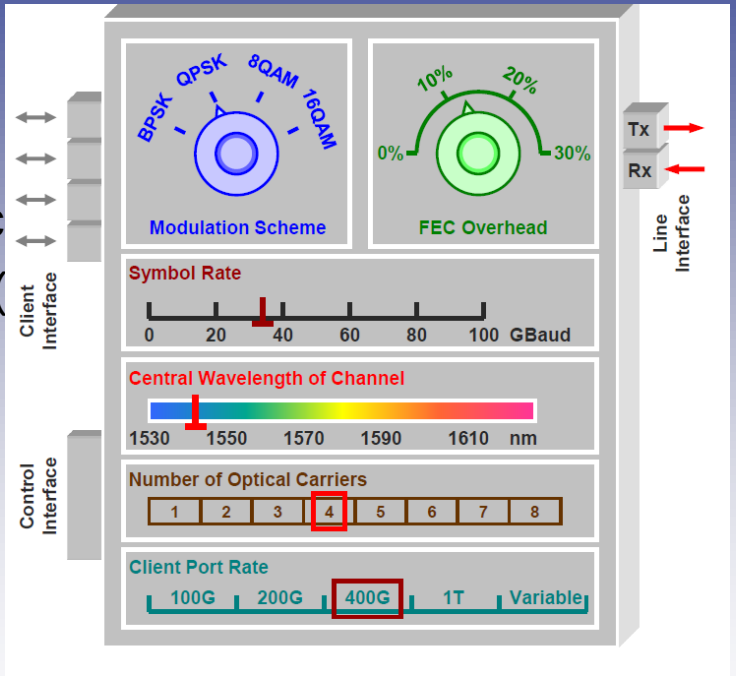
- This figure shows one prominent application for FlexEthernet
  - This is a sub rate example
  - One possibility is using a 400GbE IEEE PMD, and sub rate at 200G to match the transport capability



Ethernet Technology Summit

Matching Transport's Flexibility

- Modern transport provides flexibility
- Clients need to scale to this flexibility
  - Trans-Pacific link maybe 50 Gb/s but c
  - Router must be told to only send 50G (
- Need a simple method to adjust client transport!



Source: Keynote_Layer0vsLayer1_SDN_Wellbrock.pdf @ POTE 2013 (Verizon; Glen Wellbrock)

Ethernet Technology Summit

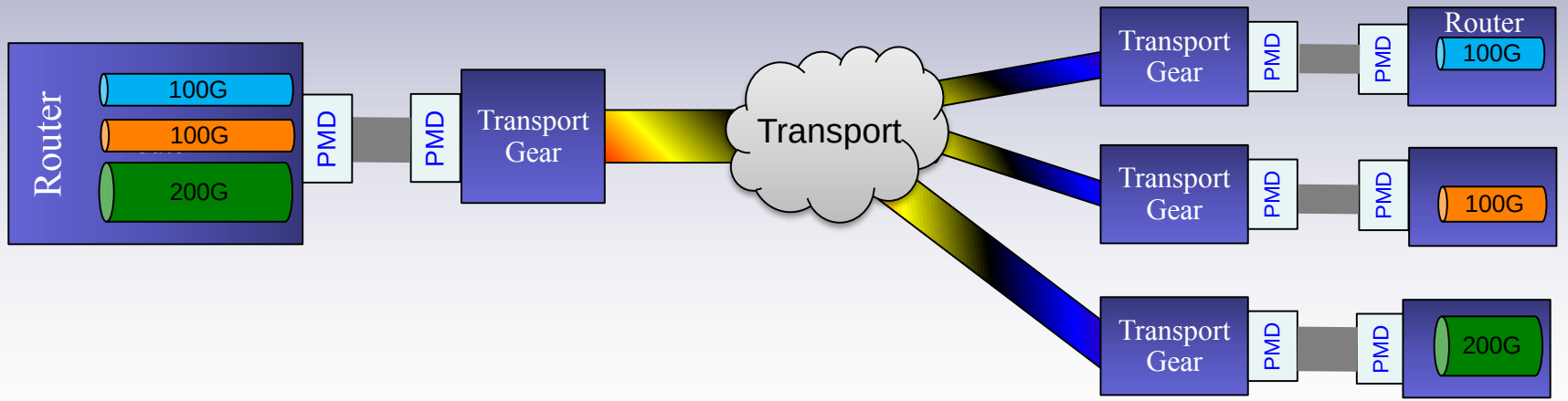
Channelized Interfaces

- Desire to support simple transport of n x Ethernet streams across a faster interface
- One example is 10x10GbE across a 4x25G interface
- A future example would be 16x25GbE across an 8x50G interface
  - Electrical or optical
- Other options are mixing speeds across a single interface

Ethernet Technology Summit

Channelized Interfaces

- One example of mixing of speeds across a transport infrastructure



Ethernet Technology Summit

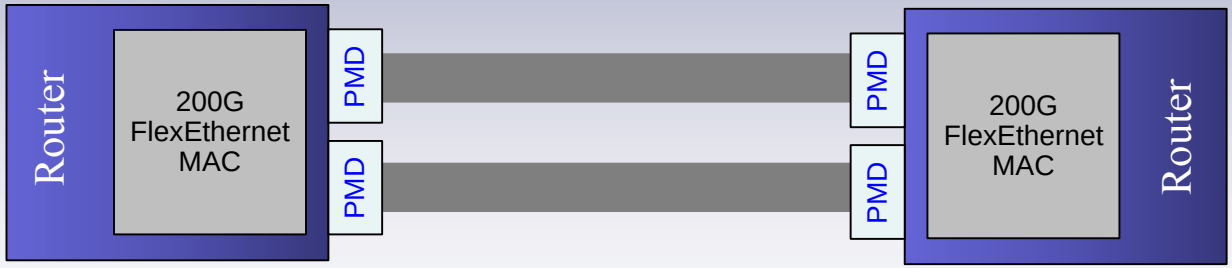
Super Rate FlexEthernet

- Link Aggregation (LAG) is often used to bundle multiple slower speed interfaces into a faster virtual interface
- LAG has a drawback of uneven distribution of flows in some cases
- There is an industry desire for an “improved” LAG
- FlexEthernet can support aggregating multiple interfaces at a physical layer, no uneven distribution of flows

Ethernet Technology Summit

Super Rate Example

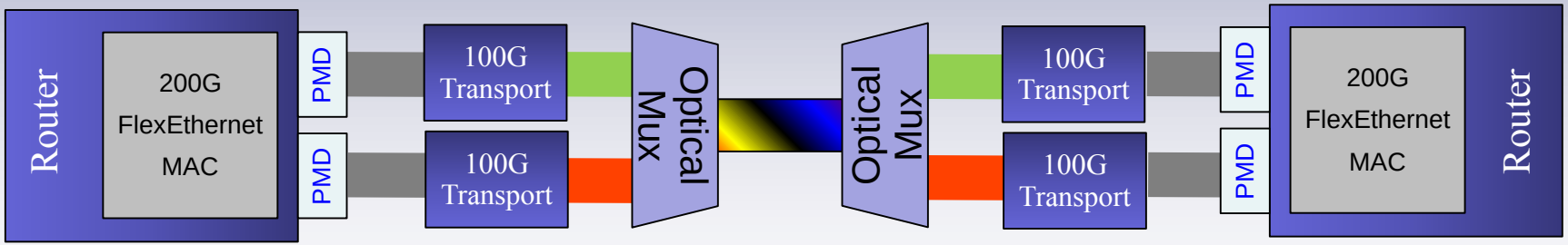
- With FlexEthernet, a 200G router to router connection could be supported, without an IEEE standard
- In the future 800G, 500G etc. can all be supported



Ethernet Technology Summit

Transport Super Rate Example

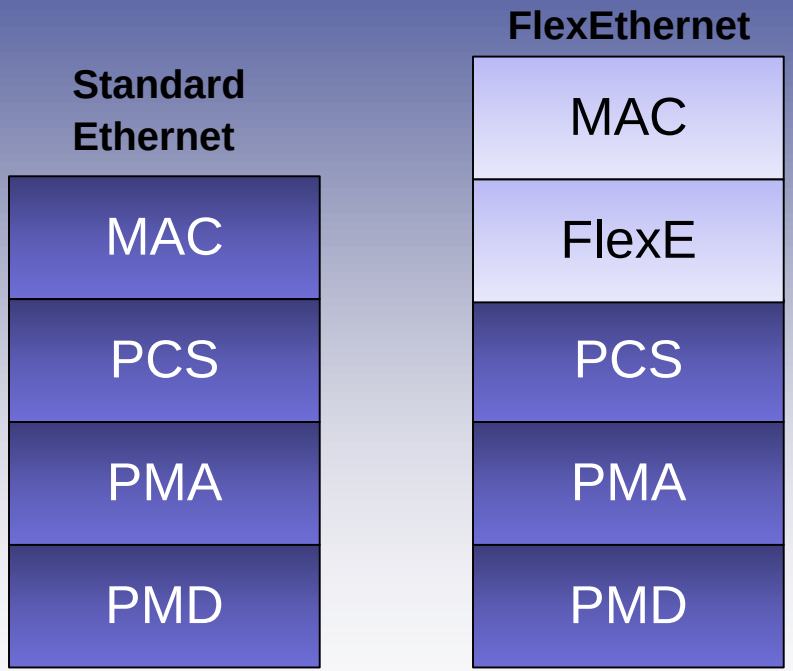
- A 200G router to router connection through current transport equipment can be supported



Ethernet Technology Summit

What is FlexEthernet?

- MAC become variable rate
- FlexEthernet adds in a shim layer
- Layer Functions:
  - MAC – Framing, addressing, error detection
  - PCS – Line coding, distribution
  - PMA – muxing, CDR
  - PMD – Physical drivers
  - FlexE – distribution, idle control



Ethernet Technology Summit

FlexEthernet Layer Requirements

- Be flexible, support variable rate interfaces
- Support sub-rate operation
- Support distribution of packets to multiple interface
- Support multiplexing of multiple streams on a single interface
- Must be transparent to interfaces that it runs across

Ethernet Technology Summit

FlexEthernet Operation

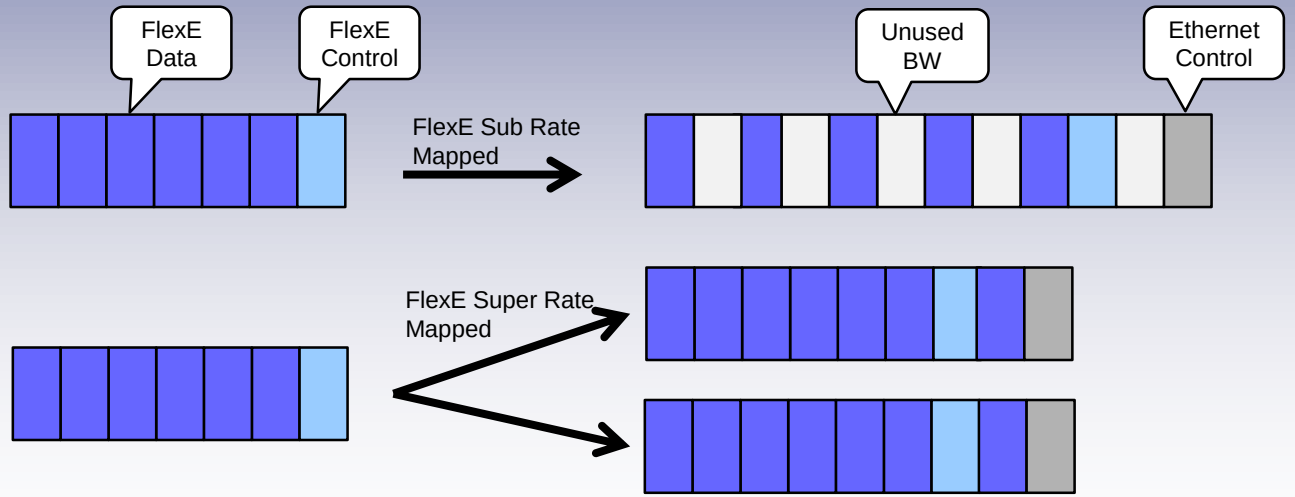
- FlexEthernet adds in periodic overhead into the data stream to allow for striping, alignment etc., via a calendar mechanism
- Overhead is based on 66b blocks structure and looks like an a sequenced ordered set
  - Sequence ordered sets also carry link fault signaling
- Ethernet idles are deleted periodically to make room for the extra overhead
- IEEE overhead such as alignment markers can also be added later in order to traverse a given PMD without impacting FlexEthernet

FlexEthernet Operation

- FlexEthernet adds in periodic overhead into the data stream to allow for striping, alignment etc., via a calendar mechanism
- Overhead is based on 66b blocks structure and looks like an a sequenced ordered set
  - Sequence ordered sets also carry link fault signaling
- Ethernet idles are deleted periodically to make room for the extra overhead
- IEEE overhead such as alignment markers can also be added later in order to traverse a given PMD without impacting FlexEthernet

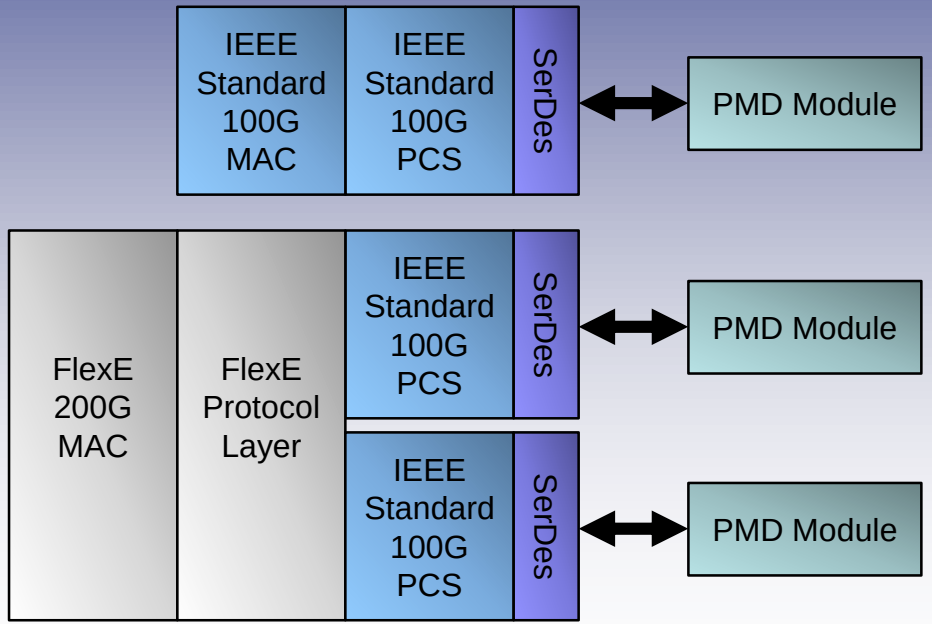
FlexEthernet's Operation

- FlexEthernet overhead is added in, and coexists with the standard Ethernet overhead



MAC/PCS Impacts

You can re-use standard IEEE PCS implementations, and add on the FlexEthernet protocol and the FlexEthernet MAC



Ethernet Technology Summit

PMD Impacts

- **The goal is to use IEEE PMDs as is for FlexEthernet**
  - For example a 100GBASE-LR4 or 100GBASE-SR4 optics module with accompanying PHY protocol stack can be re-used without modification
  - MSA defined PMDs should also work without modification, for example CWDM4
  - So no impact!
- **What this implies to the FlexEthernet protocol:**
  - Same per lane rate on the PMDs
  - Therefore we must delete extra idles from the MAC stream to make room for the FlexEthernet overhead in addition to the normal overhead used for a given IEEE PHY
    - Normal IEEE overhead is Alignment Markers for multi-lane interfaces
  - FlexEthernet must run transparently through the IEEE PHYs

Module photos courtesy of Finisar

Ethernet Technology Summit

Summary

- FlexEthernet won't replace IEEE Ethernet standards
- It will supplement and fill in the gaps
  - Faster speeds of multiple interfaces instead of LAG
  - Nx transport of Ethernet
  - Enables efficient matching of transport rates
- FlexEthernet is being standardized in the OIF
  - Proposed schedule is to have an agreement in Q2 of 2016

Thanks for your attention!