Challenges in Higher Speed Ethernet Solutions

David Ofelt- Juniper Networks
2020-01-25

Outline

- Introduction
- History
  - 400GbE
- Possible Futures
  - 800GbE (0.8TbE?)
  - 1.6TbE
- “Conclusions”

Introduction

- This talk focuses on classic LAN/WAN Ethernet applications
  - Some of the discussion may change for fibre and infrastructure heavy applications like DCI, Metro, and Long Haul
  - The new Ethernet coherent DCI optics (400ZR/802.3cw) make discussions more complicated since they live in multiple worlds

ITU IEEE 802

NOKIA

Note on IEEE CSD

- IEEE starts projects only if they meet certain criteria
- The main ones are:
  - Technical Feasibility
  - Economic Feasibility
  - Broad Market Potential
- It is important always to keep these in mind

400GbE History (IEEE 802.3bs)

- Carriers were dominant when we started the project
  - Hyperscale companies were dominant when the project finished
- A logic-only project was initially proposed to bond 4x100GbE
  - Things snowballed- added 100Gb/s/λ optics & 50Gb/s electricals
- New rate drove new technology that backfilled slower rates
  - Provided building blocks to commoditize 50GbE and 100GbE (802.3cd)

ITU IEEE 802

NOKIA

400GbE History Cont

- Deployment of 400GbE has been slower than expected
  - Historically – carriers wanted new rate despite power/cost/density hit
  - Hyperscale customers (currently) want:
    - radix over single-pipe rate
    - no cost/power/density penalty for faster rates
    - port density
    - efficient or long-lived optical format
  - Result:
    - 400GbE deployment skipped the lower-density form factors and waited for QSFP-DD/OSFP with 50Gb/s electrical lanes and 100Gb/s optical lanes

ITU IEEE 802

NOKIA

800GbE

- 100Gb/s electricals actively being worked on in OIF/802.3ck
  - These allow 800Gb/s modules to happen
  - Both OSFP & QSFP-DD will support 8x100Gb/s electricals
- Primary applications are 2x400GbE, 4x200GbE, and 8x100GbE
- No real ask from end-users yet for 800GbE (ignoring 800ZR topic)
  - Just look to the 400GbE rollout
  - Some end users want a larger than 2x jump before they upgrade
- Currently no PMD technology that is better gluing together two 400GbE
- Probably implementable in an FPGA – important for tester vendors

ITU IEEE 802

NOKIA

800GbE

- An “800GbE” will happen because it is free- not because it is wanted
  - Host ASICs likely to provide it
  - FlexE or 2x400GbE PCSs
    - Ex: https://www.morethanip.com/800g-ethernet-ip/
  - Will be useful for some end users
- Has been a suggestion that we do a logic-only IEEE project
  - Not sure it has broad market potential
  - Guaranteed to snowball ☺

1.6TbE

- Big enough step over 400GbE to be interesting for some
- Probably Technical Feasible
- Economic Feasibility and Broad Market Potential likely not good
  - Too hot, costly, and low density
- Electrical interface would currently be 16 lanes with 100Gb/s SERDES
  - COBO or CFP/8 or a new module would be needed
  - Or maybe in-package optics
- Needs 200Gb/s SERDES to fit in existing popular form factors
- FPGA Implementation will be exciting – so test equipment availability risky

ITU IEEE 802

NOKIA

1.6TbE

- Optics
  - 16 lanes of 100Gb/s/λ?
  - 8 lanes of 200Gb/s/λ?
    - Unclear there is good technology for this other than coherent
    - If you are doing coherent – just do 400Gb/s
  - 4 lanes of 400Gb/s/λ?
    - 4x400ZR(lite) - need to drop the power and physical volume by 4x over today

”Conclusions”

- Unclear what the best path forward is
- Time to start talking about it
- Community needs to be creative on electrical and optical futures
- Need a constant dialog with the end users
  - What they want now may not be what they want when they find out what they are going to get…

ITU IEEE 802 NOKIA

Thank You!

IEEE 802

NOKIA

1.0TbE

- The 2x steps in 25GbE->50GbE->100GbE->200GbE->400GbE are working well and fit the popular modules
  - Ignoring module constraints- we'll still need to support older rates
- For 100GbE- the electrical interface was 10x10Gb/s
  - Least common multiple needed to include 1,2,4,5,10 == 20
  - The factor of 5 was awkward – avoid small primes!
  - 20/20 Hindsight: might have been better to do 8x12.5Gb/s