Portfolio of short-reach link types, retimed and unretimed electrical interfaces

Piers Dawe IPtronics
Jonathan King Finisar

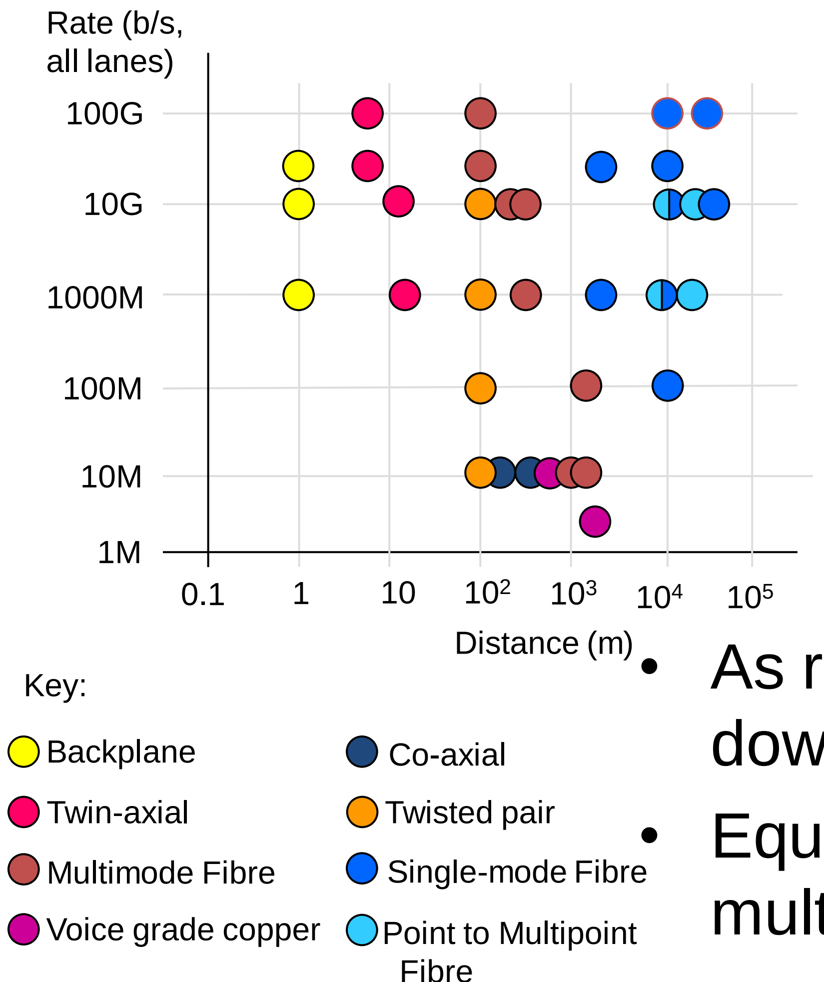
Supporters

- Robert Coenen — Reflex Photonics
- Frank Flens — Finisar

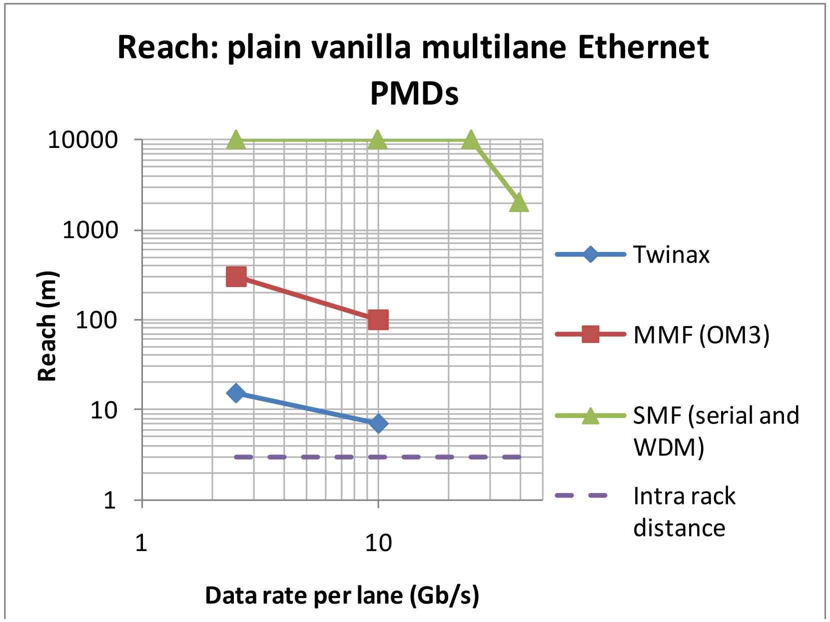
Contents

- Reach vs. rate
- Unretimed rate vs. date
- What's different
- Link types
- Unretimed and retimed electrical interface
- Parallel optics
- Reach for parallel optics
- Compatibility
- Conclusion

Introduction

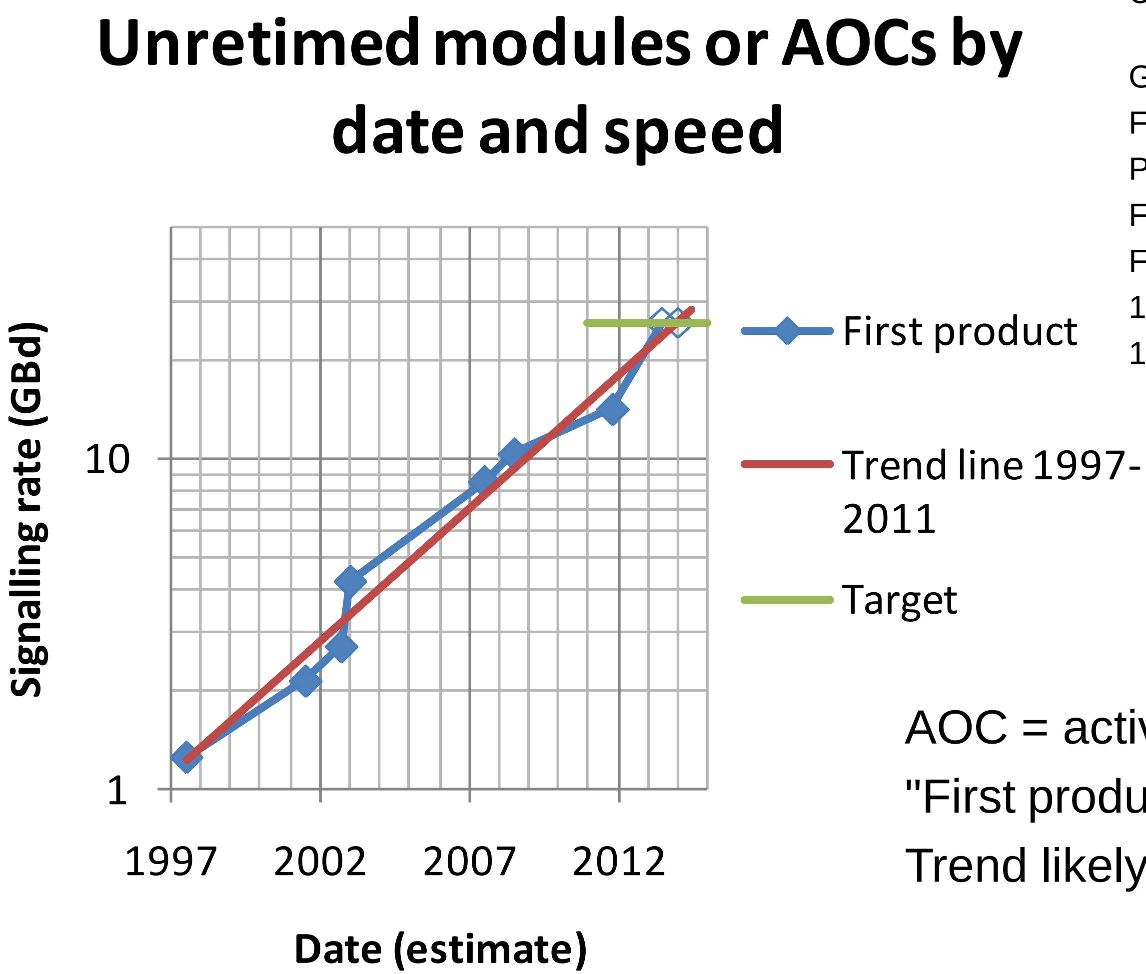


With thanks to David Law



- As rate per lane goes up, reach goes down
- Equipment rooms must deal with multiple media
- But these can plug into common sockets

Progress of unretimed interfaces



Unretimed modules or AOCs by date and speed: first product (estimates)

GBIC	1997?	1.25
FC 2G SFP	2001	2.125
POP4	2002	2.7
FC 4G SFP	2002-3	4.25
FC 8G SFP/SFP+	2007	8.5
10GE SFP+	2008	10.3125
10G QSFP+, CXP IB QDR	2008	10.3125
IB FDR	2011	14.0625
IB EDR	2013?	25.78125
100GE 4-wide	2013-4?	25.78125?

AOC = active optical cable

"First product" dates are estimates

Trend likely to slow down in future

What's different about 25 G lanes?

- ... as compared with 10G lanes

- MMF bandwidth means shorter reaches
- VCSEL speed is challenged
- More chromatic dispersion penalty
- Wider noise bandwidth – more RIN
- Reach of copper cable is reduced - more focus on truly short optical links

What's not different

- Silicon can be fast enough
- Low power, size and cost are still desired
- Optical connectors still have a cost
- Electrical connectors can be made about as good at 25 G/lane as was accepted for 10 G/lane

Possible short link types

- Six topologies are enabled by a smaller number of standardized interfaces:

	Type	Connector
1	Electrical backplane	Backplane connectors
2	Optical backplane	" "
3	Copper cable e.g. twinax	Front or back panel electrical connectors or sockets
4	Active optical cable	" " " " " "
5	Passive optical cable	Front or back panel optical connectors, mid-board O/E
6	Pluggable fibre optics	Front or back panel electrical connectors or sockets, module optical connectors, further optical connectors

- At 10G/lane, 802.3 and SFP+, QSFP+, CXP enable 1, 2 (if the waveguide is MMF), 3, 4, 5, 6
  - Standard specs for MMF are sub-optimal for short (10s of m and below) links:
    - Burden of supporting the fewer longer links,
    - May not have optimal waveguide or connector type for optical backplane
- At 25 G/lane, 802.3 and SFP+, QSFP+, CXP enable 2 (if the waveguide is MMF)
- 100GCU is working on 1 and 3
- NG100GOE Study Group should study 4, 5, 6 and could consider a more optimised support of 2
  - 4 and 5 may or may require little or no action in the Task Force

What about the unretimed electrical interface?

- The unretimed electrical interface (like GBIC, SFI, nPPI) is the desired objective, ***if*** it is viable at this speed
- Benefits
  - Compatibility across some media types, and speeds
  - Minimum power, as long as host and media channels not too demanding
  - Simpler, lower power ASIC I/O than for passive electrical cable
  - Pay as you grow with unpopulated ports
  - Compatible with optional FEC
    - E.g. for two PMD variants: minimum power/cost/size/latency and greater reach
    - No need to manage in-module CDR frequencies
- Disadvantages
  - Optical and electrical specifications should be designed together
  - On-PCB performance has to be better than either with retimed electrical interface such as XAUI, XLAUI, CAUI, or with option 5 that can use mid-board mounted transceivers
  - Requires good ASIC package performance
- Because it's harder, the unretimed interface can enter the market after the retimed, e.g. XFP then SFP+
- Because it's easier, AOCs can enter the market before pluggable modules
  - for the same link length and retimed/unretimed status

And the retimed electrical interface?

- Benefits
  - Compatibility across media types
  - Simpler, lower power ASIC I/O than for passive electrical cable
  - Relaxed IC and PCB performance requirements
  - Pay as you grow with unpopulated ports
  - Compatible with optional FEC
    - E.g. for two PMD variants: minimum power/cost/size/latency and greater reach
  - Independent analog interfaces, reduced test cost
- Disadvantages
  - Need to plan for all signalling rates (e.g. 10G, 25G, 25G+FEC)
  - Should be designed together at the same time as or after the unretimed interface
  - Example CDR power consumption might be 150 to 250 mW/lane each way depending on process
    - or 1.2 W to 2 W per module or cable end for 4 lanes

- Is the retimed electrical interface viable at 25G/lane?
  - It's "only" 4 lanes each way not 10 lanes each way
  - Conventional NRZ is suitable for both electrical and optical parts of the link, no need for recoding
  - 25G/lane CDRs are beginning to become available
- Conclusion: yes. Should be a subset of the unretimed electrical interface

- Is there merit in half-retimed? If so, which end?
  - For study

Considerations for unretimed electrical interface at 25 Gb/lane

- Challenges
  - Frequency-dependent PCB loss has to be compensated somewhere
  - ASIC package, PCB and fibre impairments come out of the same jitter budget
  - Noise, electrical crosstalk and reflections
  - VCSEL speed
- Solutions
  - PCB loss compensation
    - There is a range from simple to sophisticated
  - Fast low noise lasers
  - An active optical cable (AOC) avoids two or more optical connectors and allows optimised setup
    - so can be faster, or lower power, or lower jitter, than pluggables for the same technology.
  - An AOC's reach is known and usually shorter than the longest supported reach for pluggable

Considerations for retimed or unretimed at 25 Gb/lane

- To think about
  - Wavelength
    - Reduce chromatic dispersion at e.g. 980 nm rather than 850 nm
  - FEC
    - Used in 10GBASE-KR (optionally), 10GEPON (always), and 40/100GE copper and backplane (optionally). Two schemes
    - Expect that 100GCU will use FEC also – could be a new scheme
    - Trade off between latency, overhead and benefit to budget.
      - The round trip time of a 30 m passive cable alone is 2 * 30 m * 5 ns/m * 100 Gb/s = 30,000 MAC bit times
    - All 25G/lane types (electrical and optical) would benefit from FEC
    - For optical; trade off between FEC, RIN, and to an extent, MMF or PCB reach
  - Equalisation of the optical signal
  - Test cost
    - Retimed interface may need less testing than unretimed
    - AOCs have fewer analog interfaces to test than pluggables

Relative capabilities

- **Maybe** technology that's good enough to deliver
  - 10G unretimed pluggable modules (e.g. SFP+ or nPPI/QSFP+) could deliver
  - 13G unretimed AOC, or
  - 14G retimed pluggable, or
  - 16G retimed AOC
- Or restated another way, a similar technology level could deliver retimed pluggable or unretimed AOC
- If retimed, the AOC could have greater reach, or lower power, than the equivalent pluggable.

Rates available now and in the near future for parallel optics

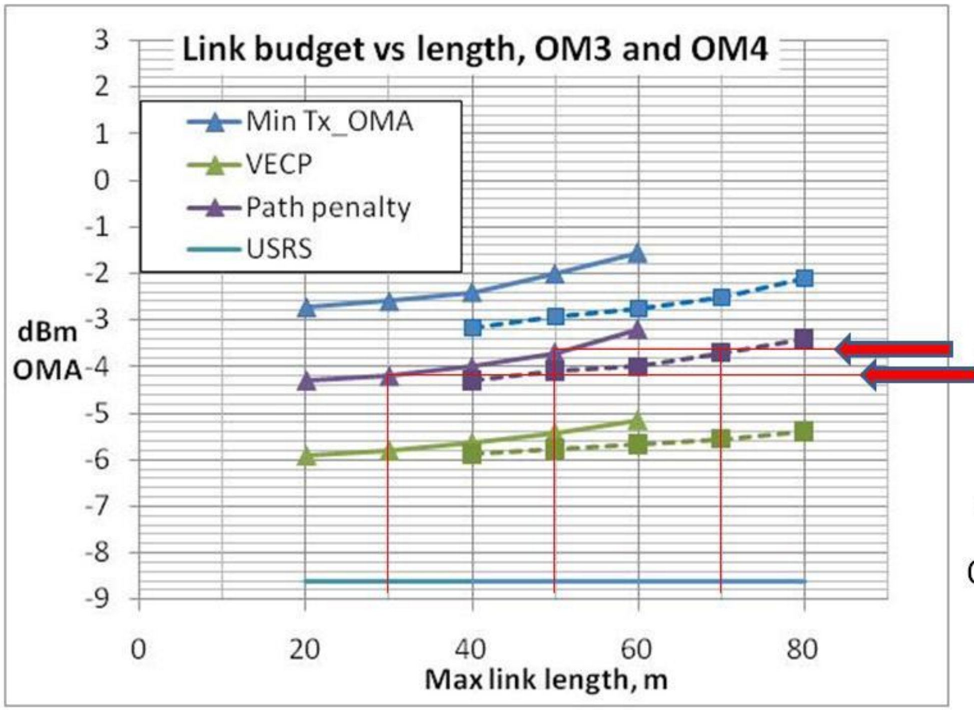
  - Generally, each generation (speed) of small parallel optics modules and AOCs has been unretimed
  - Serial speeds on MMF are introduced a little earlier than parallel
- 10 GBd InfiniBand QDR has been available since 2008
- 10.3125 GBd 40GBASE-SR4 and AOCs have been available since 2010
- 14.0625 GBd InfiniBand FDR AOCs "will be generally available sometime in the second half of this year"
- Target is 25.78125 GBd
  - Or slightly higher with FEC
- So only a factor of 1.85 in speed to go!
- According to the trend on slide 5, that might happen in 2014-5
- Expect unretimed pluggables to follow two or three years behind

What reach is appropriate?

- This is all for data centres/equipment rooms/supercomputers/central offices
  - No Broad Market Potential for 100GE in campus wiring in timeframe of this project
  - This presentation does not address the longer reach, SMF, "big module"
- Have to reserve some of the budget for electrical connector and host PCB
- Do not push fibre length to the max
  - Do not repeat the 300 m/OM3/SFP+ difficulties of 10GBASE-SR, learn from the interest in 10G "USR"
- There are many more short links than long links
- All but intra-rack links likely to be optical
- Therefore optimise for cost of the short links (3 m to 30 m)
  - More survey information in this area would be very valuable
- Most of these short links can be AOCs
  - There should still be a cost-optimised pluggable spec
  - Is there Broad Market Potential for a separate, longer reach 4x MMF spec?
  - Would this be entirely different or interoperable?

Fully retimed single lane 28 GBd model

Retimed model results



OM3 solid lines

OM4 dashed lines

Extract from T11/11-241v0 Jonathan King and Jim Tatum, June 2011

- A 'simple' retimed module with expected Tx and RX parameter values may allow 30 – 50 metres on OM3, 50 -70 metres on OM4
  - (for 4.4 to 5 dB max path penalty)
- Longer reach if Tx parameters improve: Tx Rise/fall time, RIN, most critical

7

Connector loss allowance: 1.5 dB for OM3, 1 dB for OM4. RIN_xOMA -130 dB/Hz

Unretimed, 25.78125 GBd

Host Tx jitter	nPPI	16GFC
RIN_xOMA= -132.82 dB/Hz		
OM2	10	-
OM3	25	-
OM4	31	-
RIN_xOMA= -130 dB/Hz		
Any fibre	Too much TJ at TP4	
RIN_xOMA= -130 dB/Hz and KR FEC		
OM2	2	5
OM3	66	10
OM4	89	11

- Predicted reach in metres – very preliminary, subject to change
- Same model as previous slide, but with host jitter
- 18 ps rise time
- Criteria are <4.4 dB link penalties at eye centre and <0.71 UI TJ at TP4
  - Jitter limited without FEC
- Connector loss 1.5 dB for all fibre types
- nPPI jitter seems too optimistic at 25G – jitter in UI may be more similar to 16GFC
- Ethernet line rate is 8% slower than 32GFC – makes a worthwhile difference
- This represents pluggable modules. AOCs would tolerate a little more jitter
- This FEC has a coding gain of 1.1 optical dB. Stronger codes are available.

Compatibility

- Seek a very high level of compatibility between retimed and unretimed
  - Compatible voltages, reflection specs and so on
    - CPPI-4 signals wholly compliant with CAUI-4 – minimal mode setting needed for interoperability, can use retimed module in unretimed slot
    - E.g. CPPI-4 reflection spec should be within CAUI-4 reflection spec
    - Because there is more design freedom for CAUI-4, design CPPI-4 first
  - Probably same form factor e.g. QSFP+ whether retimed or not
- Backward compatibility with XLPPI (nPPI for 40GBASE-SR) and XLAUI
- Compatibility with 40GBASE-CR4 and 100GBASE-CR4
  - Host reads the module/cable registers and sets its mode accordingly
- Compatibility with InfiniBand EDR gains economies of scale

  - See "Compatibility of Different Port Types at a Big IC", http://ieee802.org/3/ba/public/jul08/dawe_03_0708.pdf for much more detail in the context of 10G lanes

Conclusion

- A retimed "CAUI-4" would be viable for short optical links until faster lasers came along
  - but would have a longer life for optical links that contain a CDR anyway, e.g. because they contain DFE or use QPSK, or multiplex to 50 or 100 Gb/s/lane
- An unretimed "CPPI-4" would be viable in the time horizon of this project
- Both retimed and unretimed electrical interfaces should be part of this project
- Focus clearly on low cost for the majority of links
  - E.g. 3 m to 30 m
  - Plan for a retimed and unretimed future
  - Do not repeat the 300 m 10GBASE-SR mistake