How Much Optics Does AI Need?

Rump Session

OFC 2024
March 26, 2024
Chris Cole, Quintessent

Optical Computing: Real or Sci-Fi?

- Energy: No savings over digital CMOS HW computation
    - ➢ Dominated by Processor to Memory data movement
- Area: 10,000 to 1,000,000 larger than digital CMOS
- Precision: ~4 bits, however INT4 not useful standalone in the Datacenter
    - Higher precision, like 8 bits, claims unsubstantiated
        - ➢ Rigorous measurement of ENOB never reported
    - A priori, the required precision of AI applications is unknown and unpredictable
    - Digital CMOS HW is programmable INTm and FPn, low-to-high precision
        - ➢ If useful, INT4 or 2 dedicated CMOS would be tiny and trivial do develop
- Sci-Fi: For many reasons, production Optical Compute HW does not and will not exist in the Datacenter, claims to the contrary not withstanding

Datacom Optical Interconnect: Real or Sci-Fi?

- Real: technical data
    - Shannon Hartley: $C = B \log_2(1 + S/N)$
        - ➢ measured full link BER curves
    - Reliability: System MTBF or FIT
        - ➢ accelerated aging data calculations
    - Multi-channel: WDM, TDM, Spatial, other
        - ➢ BER measured with all channels ON
    - Integrated: CPO, Chiplets, Wafer-scale
        - ➢ BER measured with entire system ON, over all process and temp corners
- Sci-Fi: exponential growth curves, expansive claims, cartoon system renderings
    - ➢ No statistically valid measured link BER curves
    - ➢ No system FIT data-based calculations