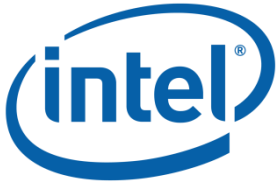
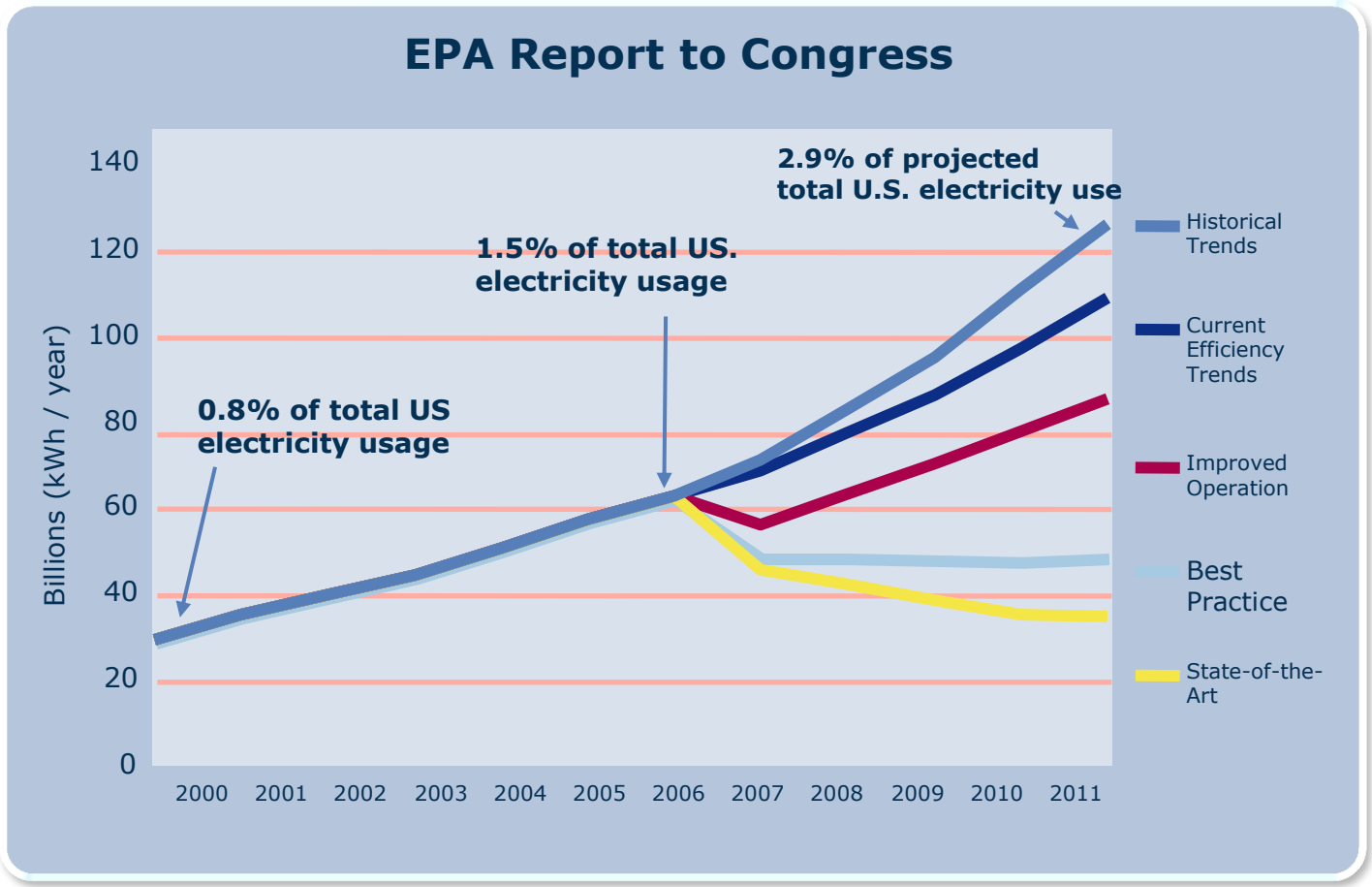


Power Efficient Bandwidth Delivery for the Data Center

Randy Mooney

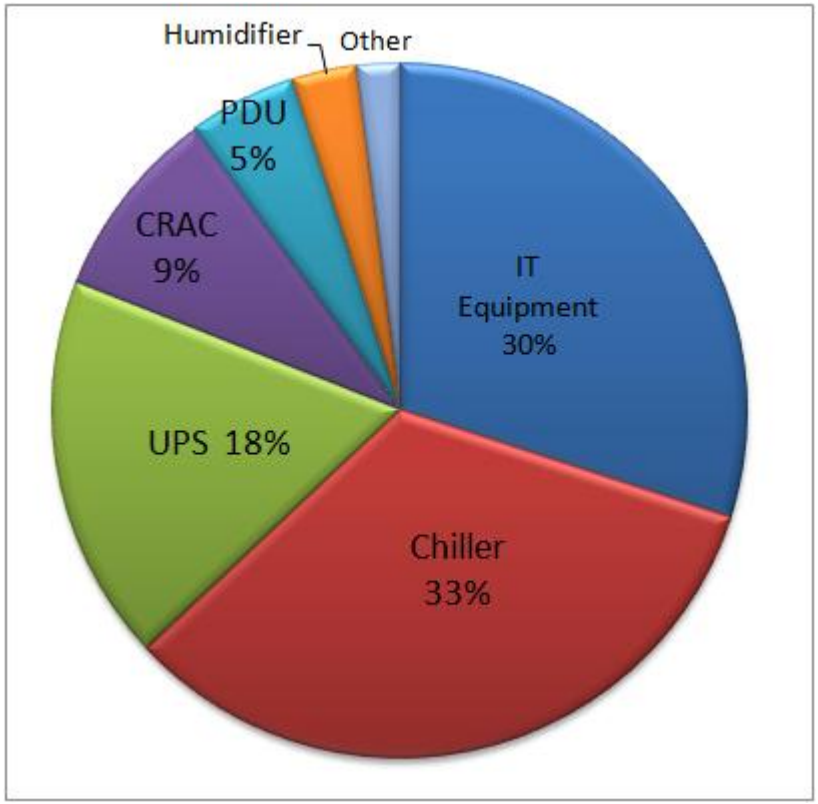
US Power Consumption from Data Centers



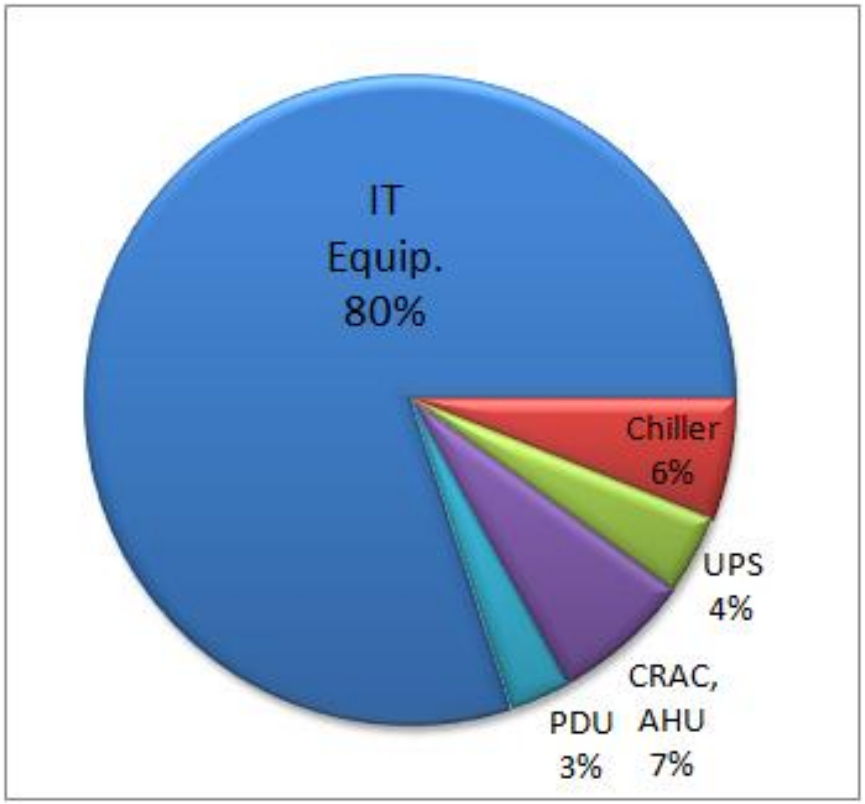
Source: EPA Report to Congress on Server and Data Center Energy Efficiency; August 2, 2007

2007 Report to Congress highlighted the potential problems from growth in Data Center demand.

Power Breakout



Circa 1990 – 2005

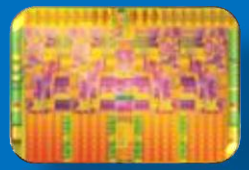


Today's New Data Centers

Concentrate future power reduction on core IT

Intel Optimization Approach

Optimized Silicon



Low voltage processors
Tailored SKUs
Efficiency features

Optimized Technologies

Power Management
Security Technologies
Solid State Drives
Advanced Networking

Software Optimization

Parallelism
Scalability
Configurations
Manageability

Optimized Systems

Optimized boards
System tuning
Rack optimization
Power tuning

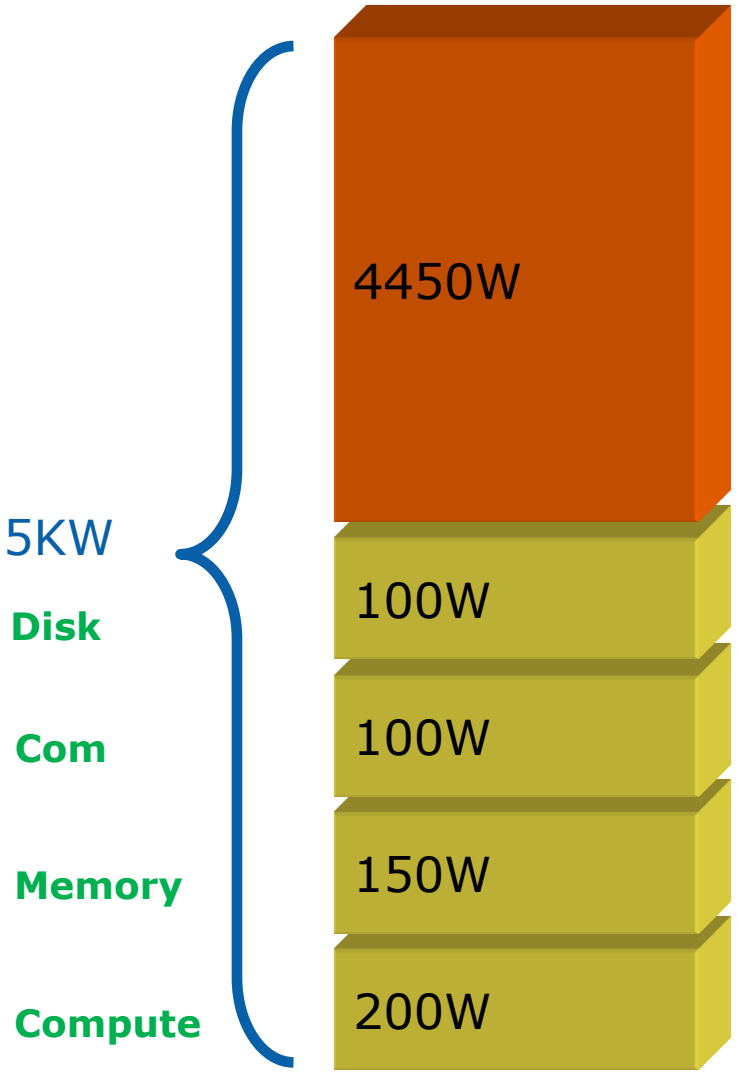
Datacenter Optimization

Floor Plan
Aisle Layout
Integration
Operating Conditions

Efficiency Losses Cascade

Building with Today's Technology

TFLOP Machine today



Decode and control
Translations
…etc
Power supply losses
Cooling…etc

10TB disk @ 1TB/disk @10W

100pJ com per FLOP

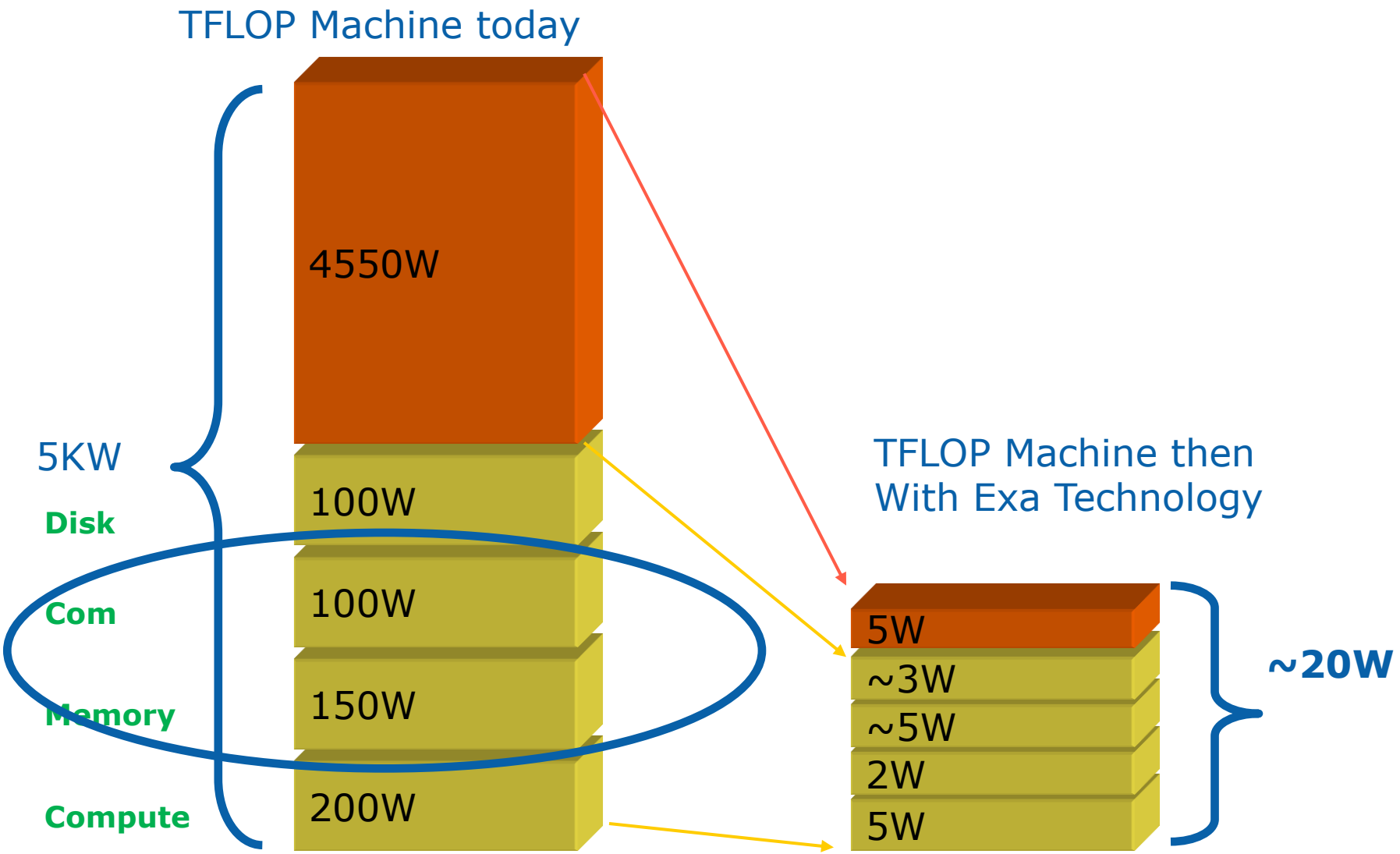
0.1B/FLOP @ 1.5nJ per Byte

200pJ per FLOP

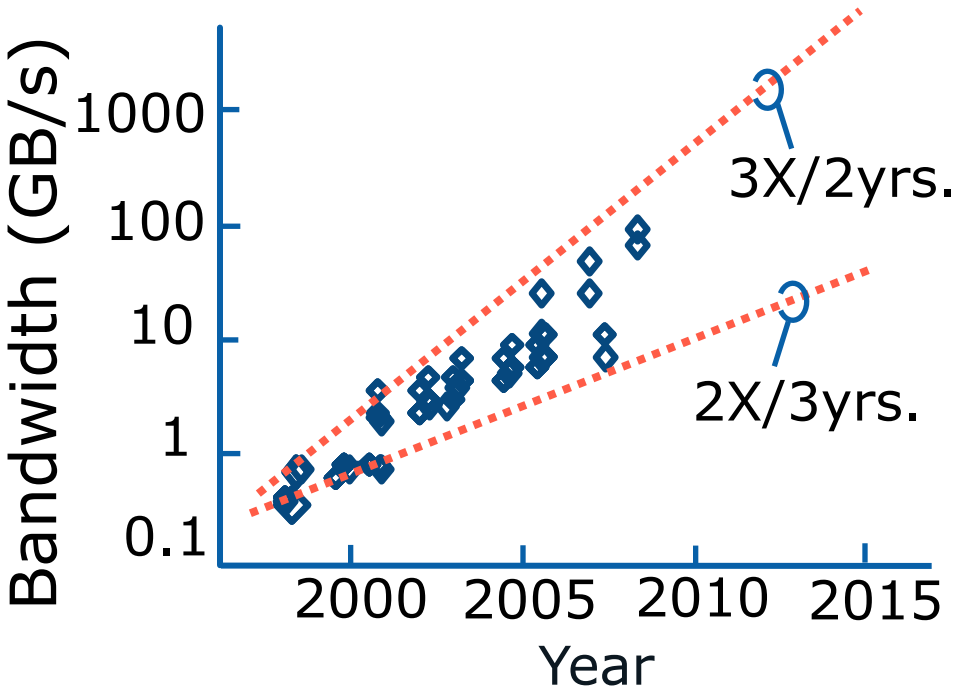
KW Tera, MW Peta, GW Exa?

intel

The Power & Energy Challenge



Microprocessor Bandwidth Trends



Bandwidth Drivers:

CPU↔Memory

CPU↔CPU

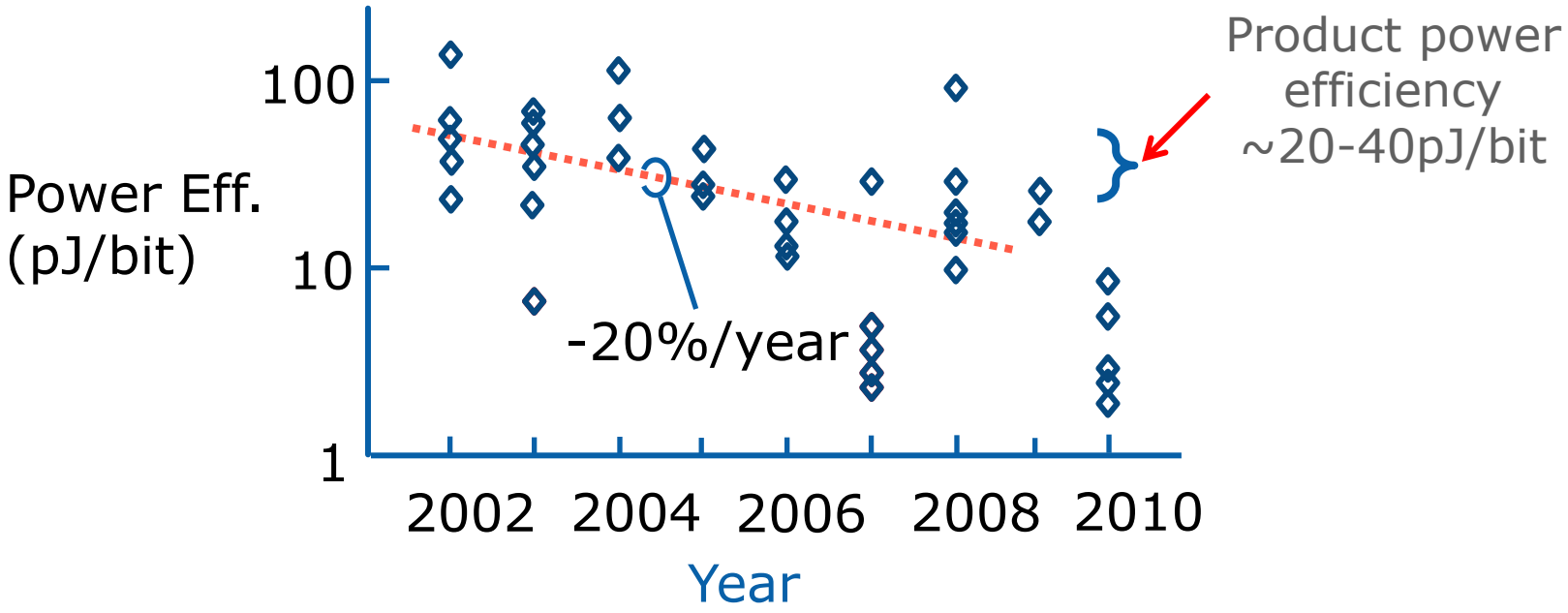
CPU↔Peripheral

CPU↔I/O bridge

Most apps <1m length

High-end microprocessors expected to need ~1TB/s by 2020

Trends in I/O Power vs. Year*

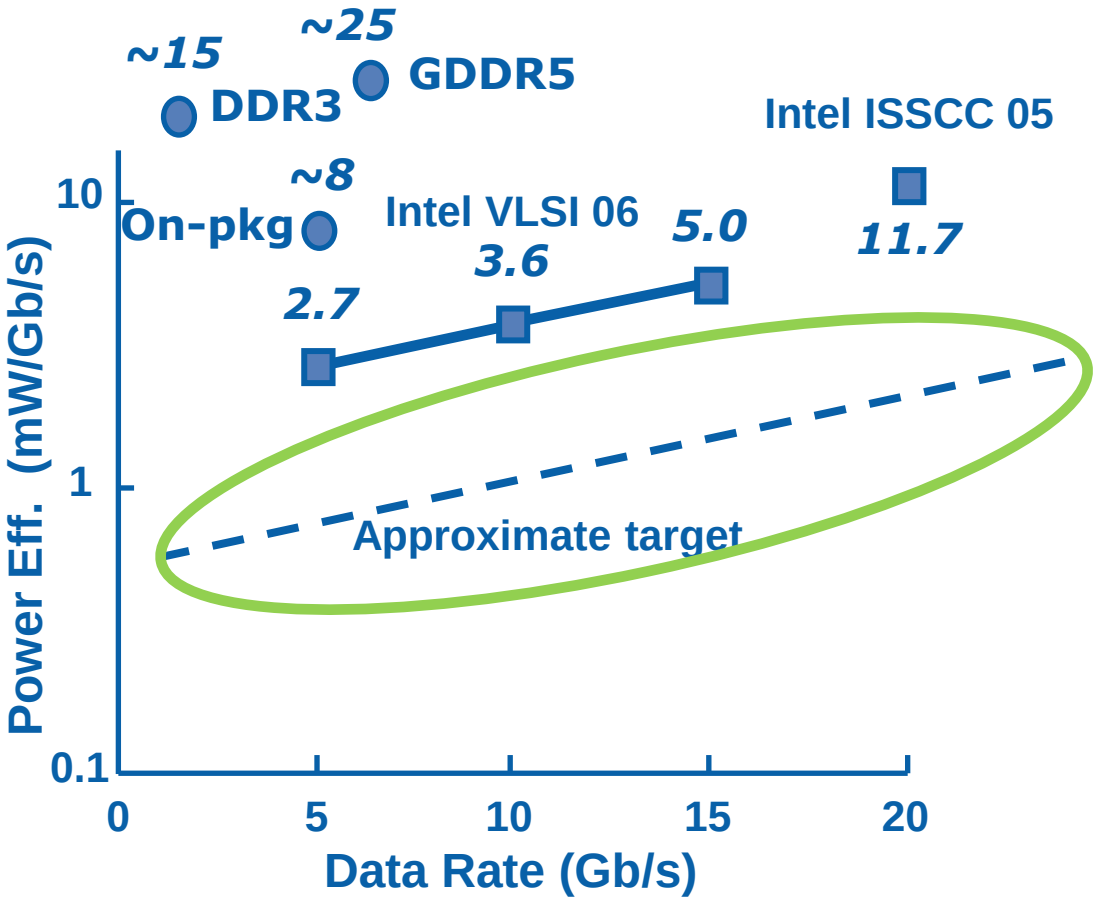


Issue: ~20% per year power reduction while bandwidth increasing ~2x every 3 years

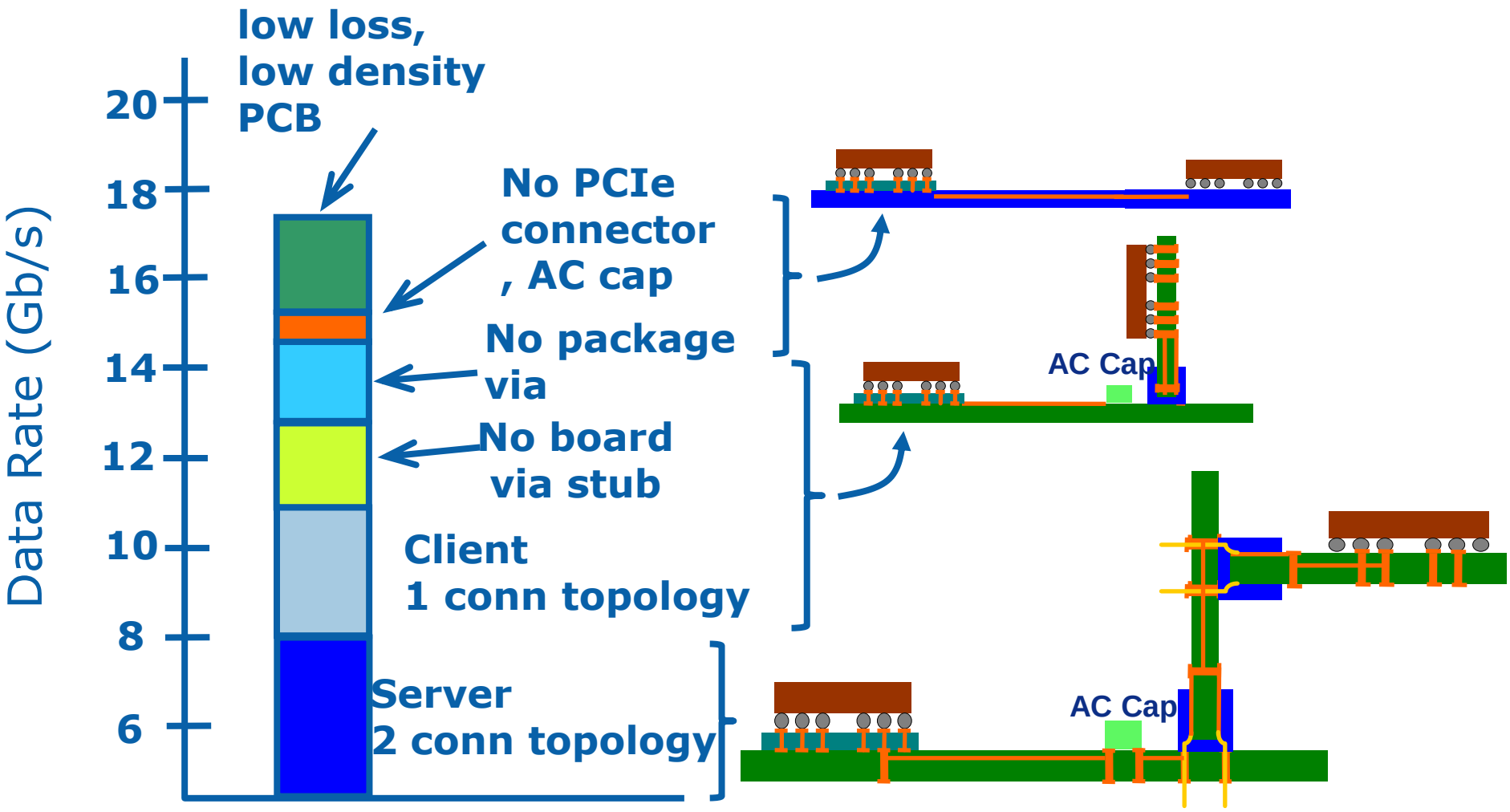
*Non-scientific sample of top-tier peer reviewed publications

Ideal Interconnect

- BW scalable across 3 platform generation minimum
- Best possible power efficiency
- Reconfigurable to fit multiple channel types
- Scalable bandwidth/power
- Fast entry/exit to/from lowest power state
- High density
- Distance solution



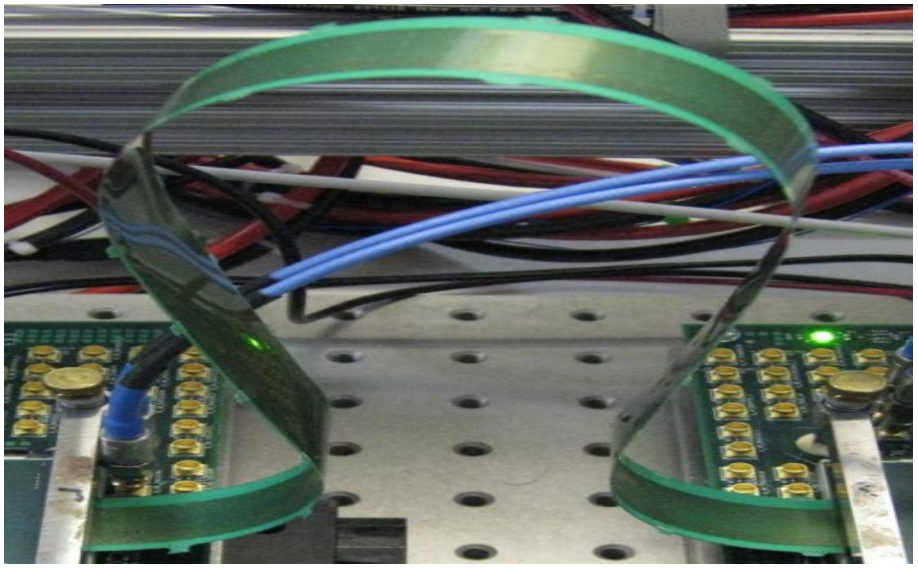
Evolutionary Interconnects: PCI-E Example



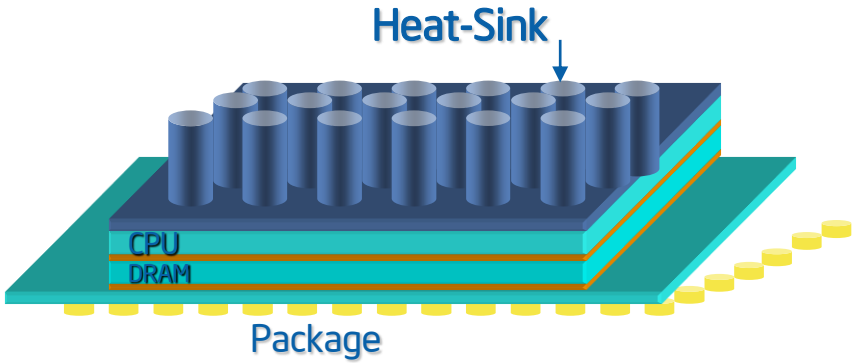
Conventional electrical interconnects nearing EOL
Using *all* evolutionary improvements *may* buy a generation
Now is the time to make a break to a scalable solution

Possible BW Solutions

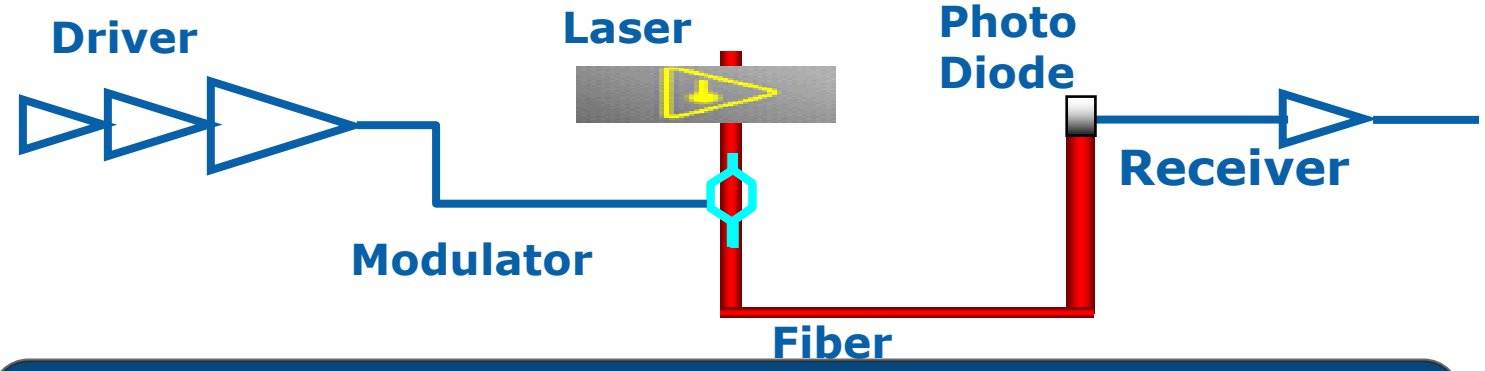
Advanced Electrical Interconnects



3D Stacking

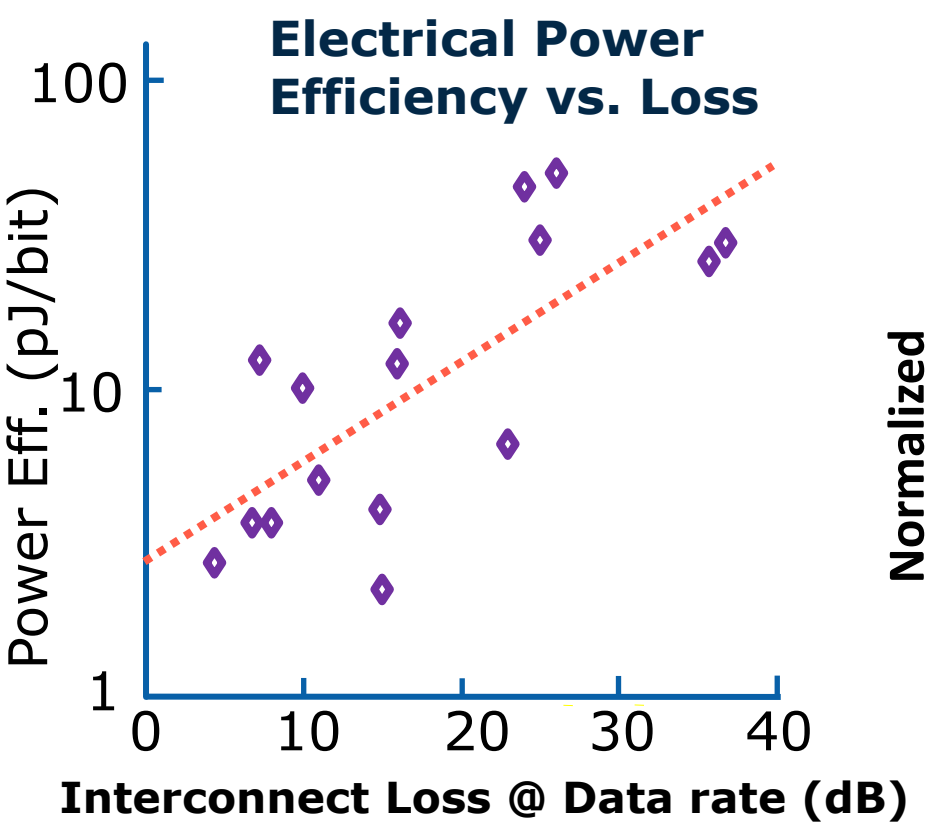


Optical Interconnects

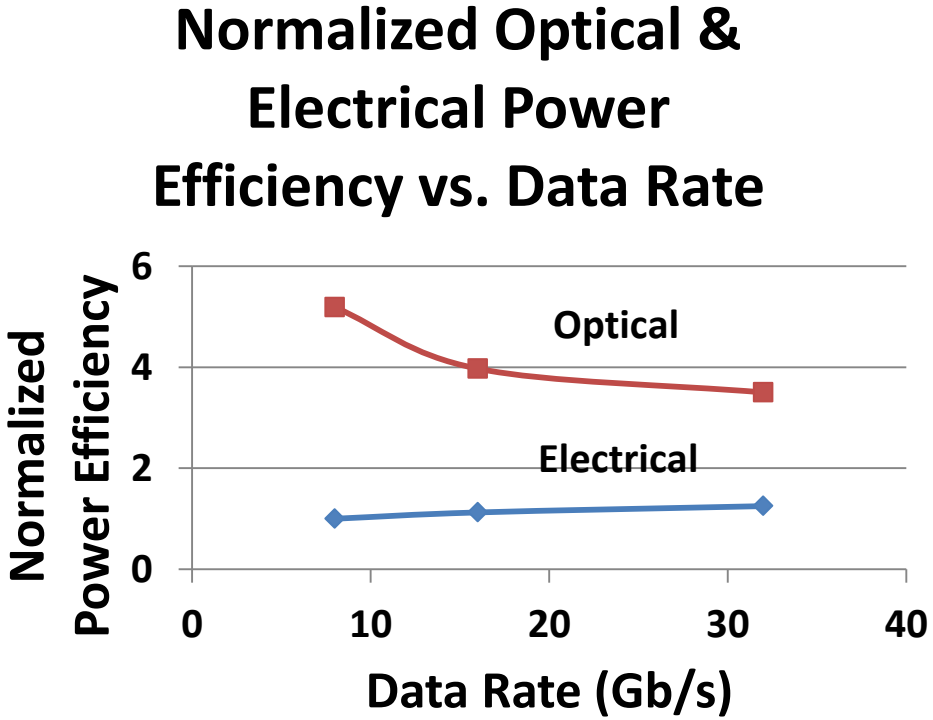


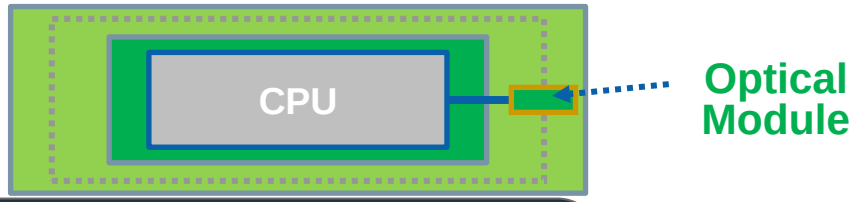
New technologies emerging
None of them solve the whole problem
Use all of these in optimal/innovative combinations

Electrical/Optical Power Comparison



(Based on transceivers reported 2006-2009 in 65-130nm CMOS)

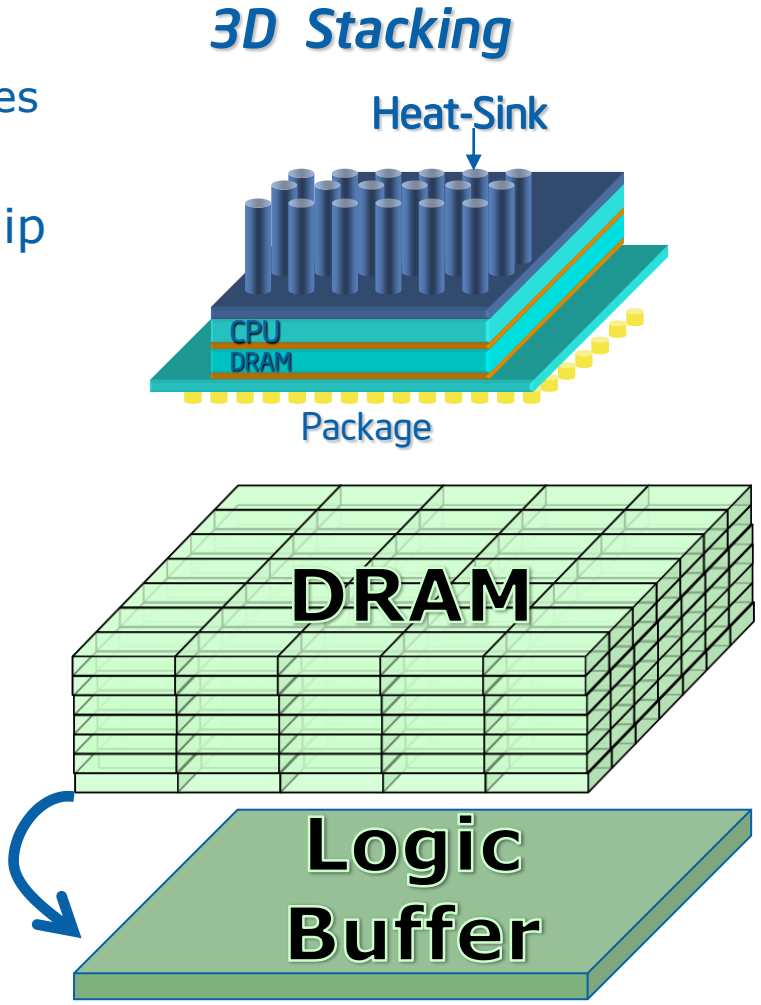




- Increased bandwidth means increasing I/O data rates
- Interconnect loss increases with data rate and distance
- Need elect interconn to optical modules , so no cross-over
- ***Moving bits across distance costs power***

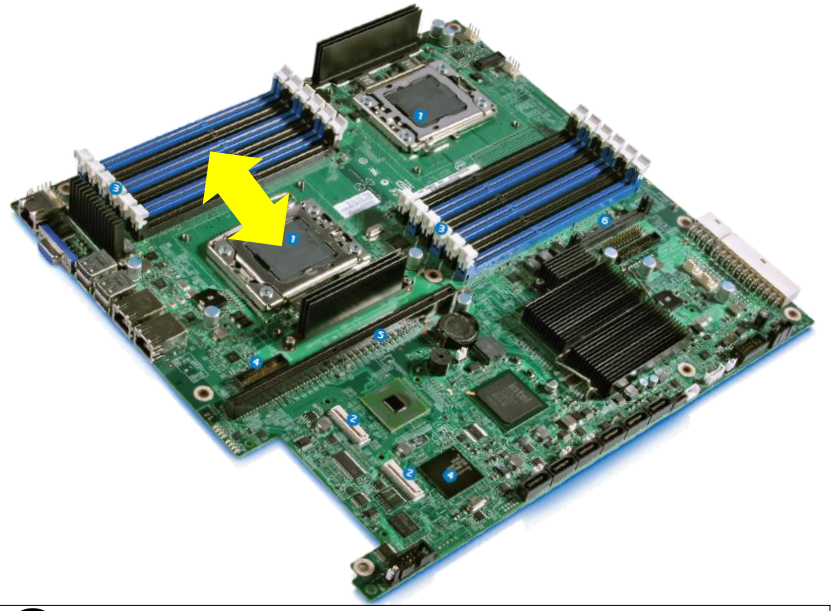
Hybrid Stacked DRAM

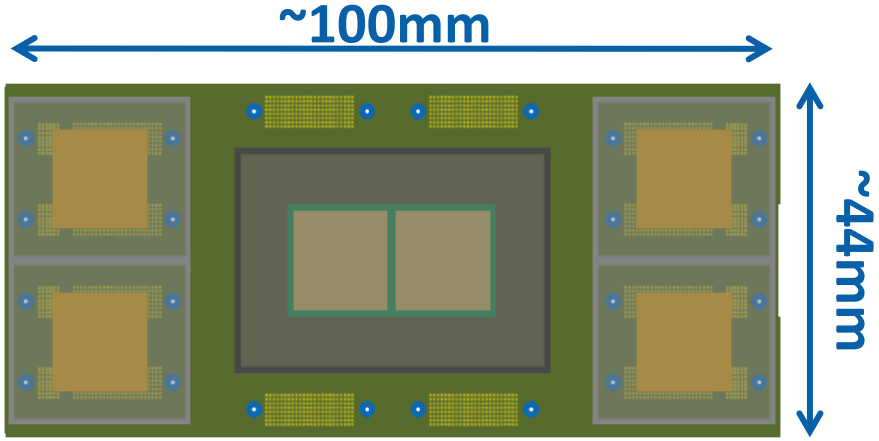
- DRAM stacked with CPU
  - Works great for low power SOC
  - Severe thermal and power delivery challenges for high power processors
- So… Stack DRAM with a dedicated logic chip
- DRAM die optimized for:
  - Memory density, static power, cost
- LOGIC die:
  - Optimized for logic density, active power, performance
  - Offload clocking, I/O, logic from DRAM
  - High BW with good power efficiency
  - Enables "smart" memory
  - Interface more appropriate for CPU
- Wide, slow interface to DRAM, serialize in logic buffer

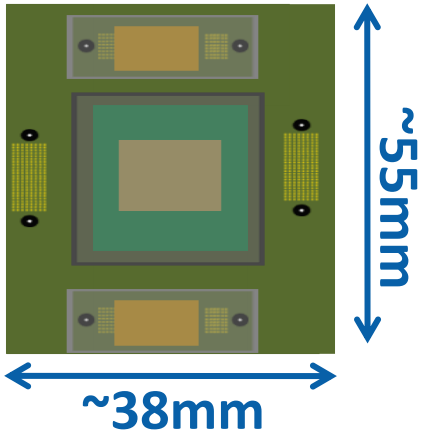


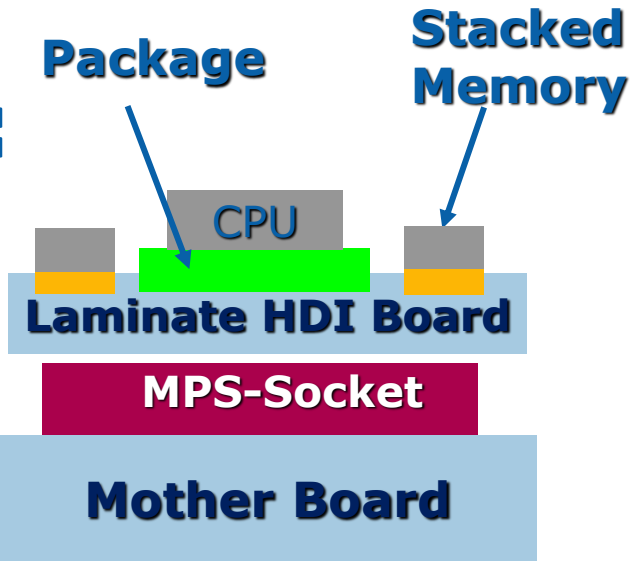
Optimal silicon partitioning enables BW

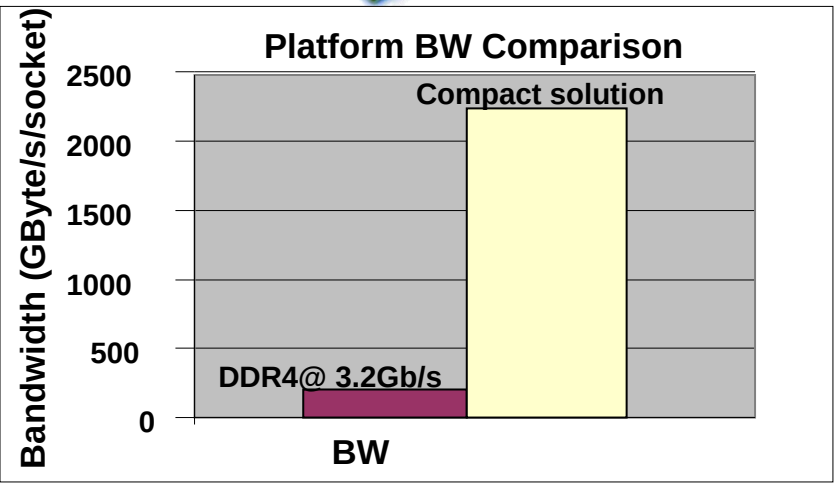
Compact the Platform







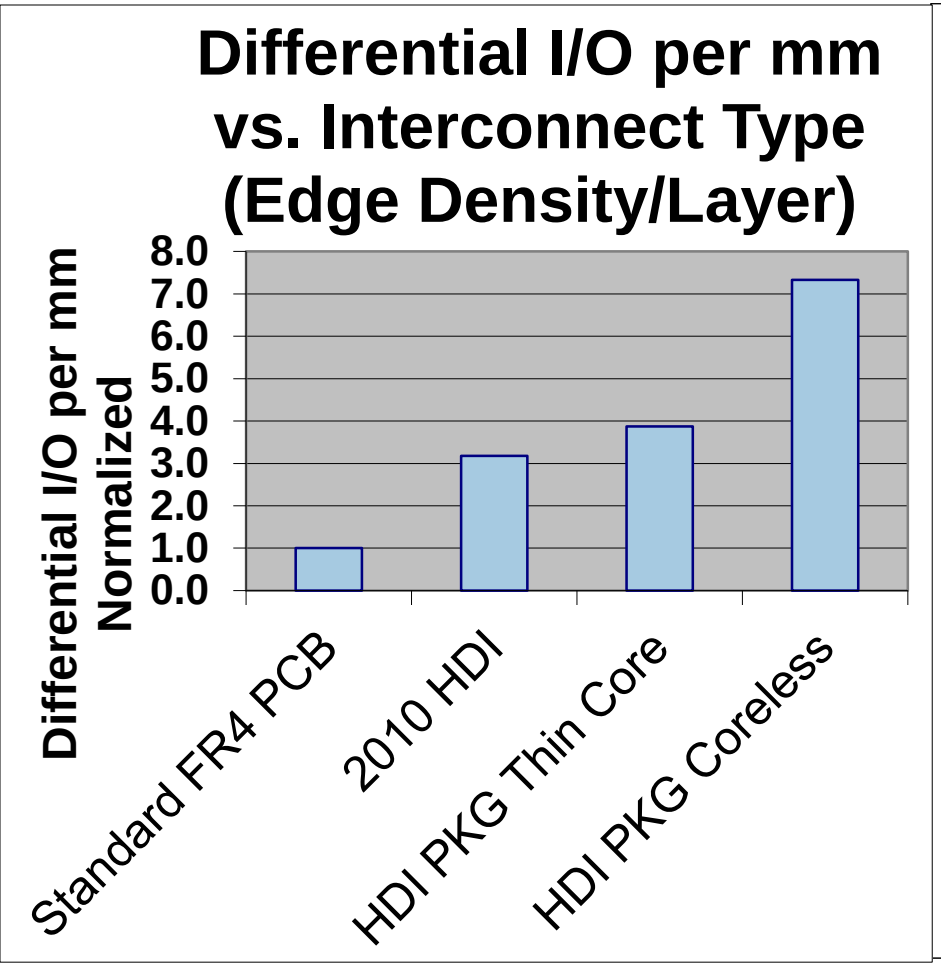


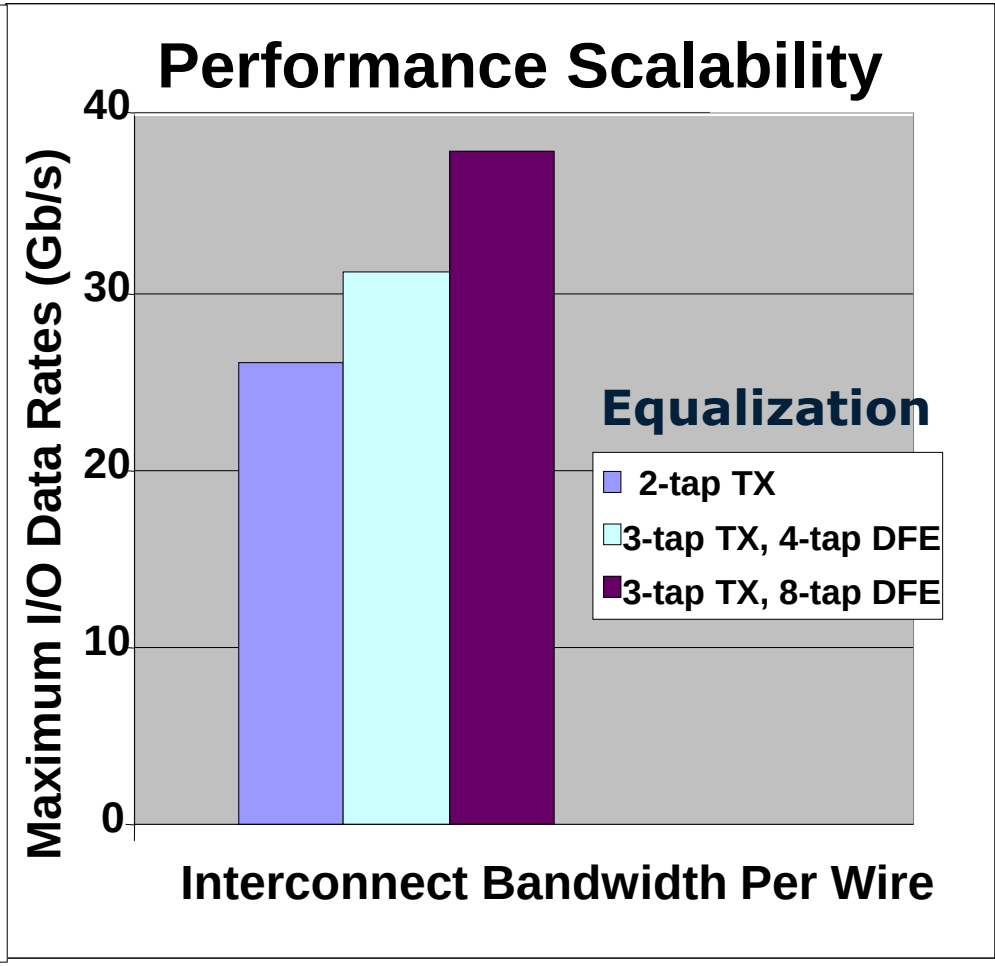


Big BW boost for compact solution

intel

Interconnect Density and Data Rate





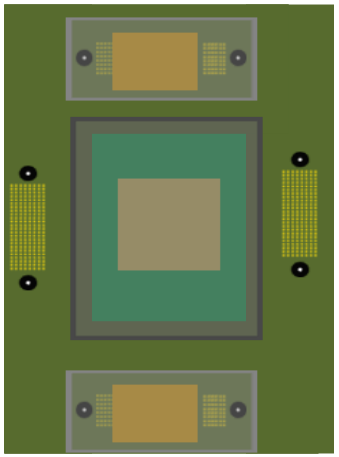
Significant density increase

Scalable pin data rate to >32Gb/s

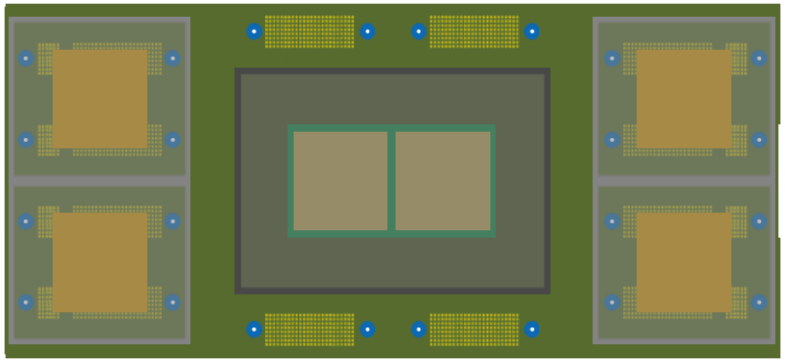
Use short, dense electrical interconnects for most cases

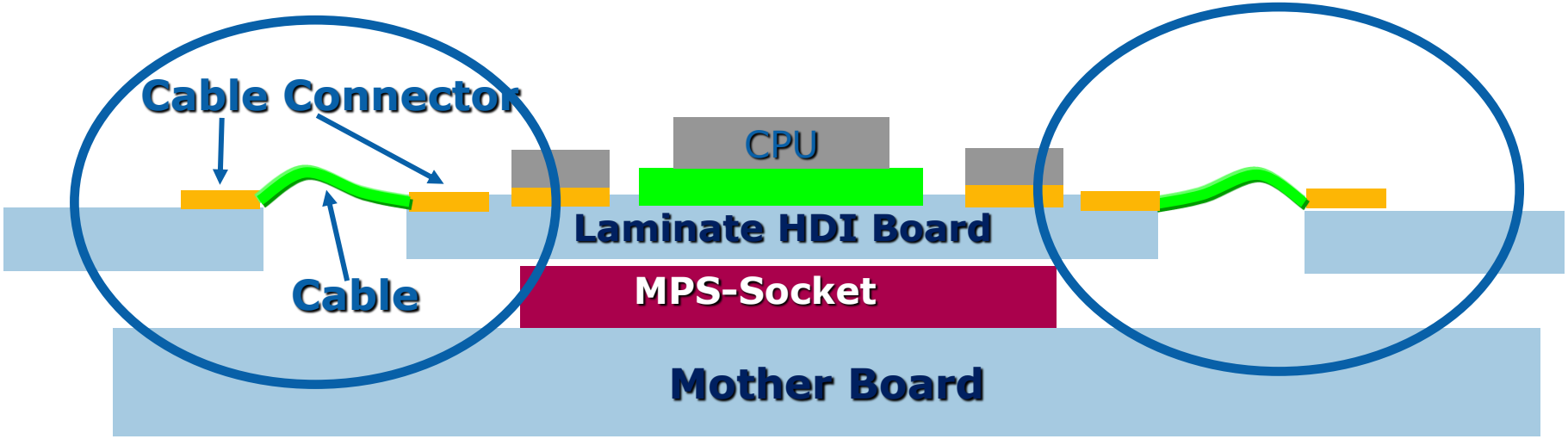
Cabled Interconnects

MS Server



HE HPC



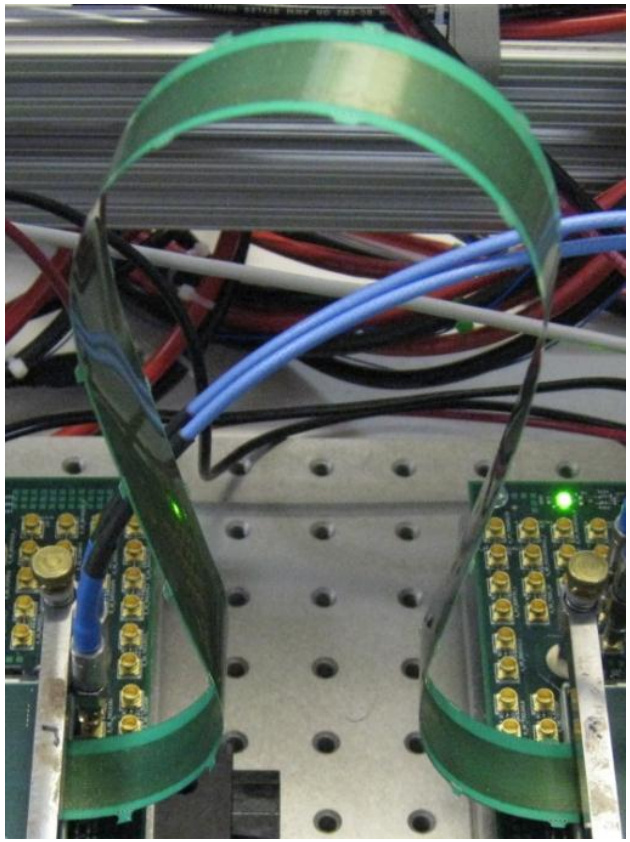


Cabled connections from CPU- avoid MB

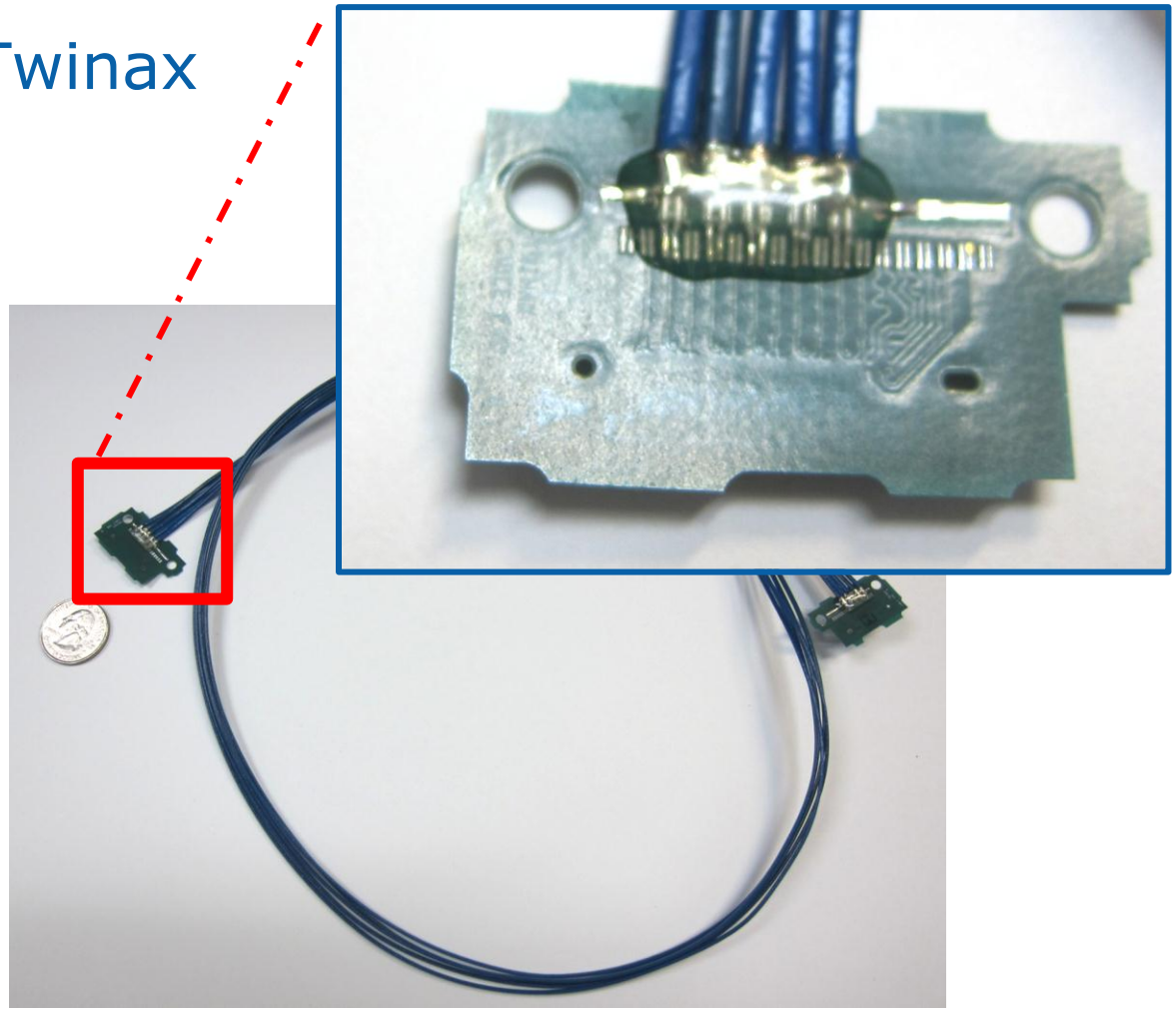
intel

Non-Traditional Interconnects

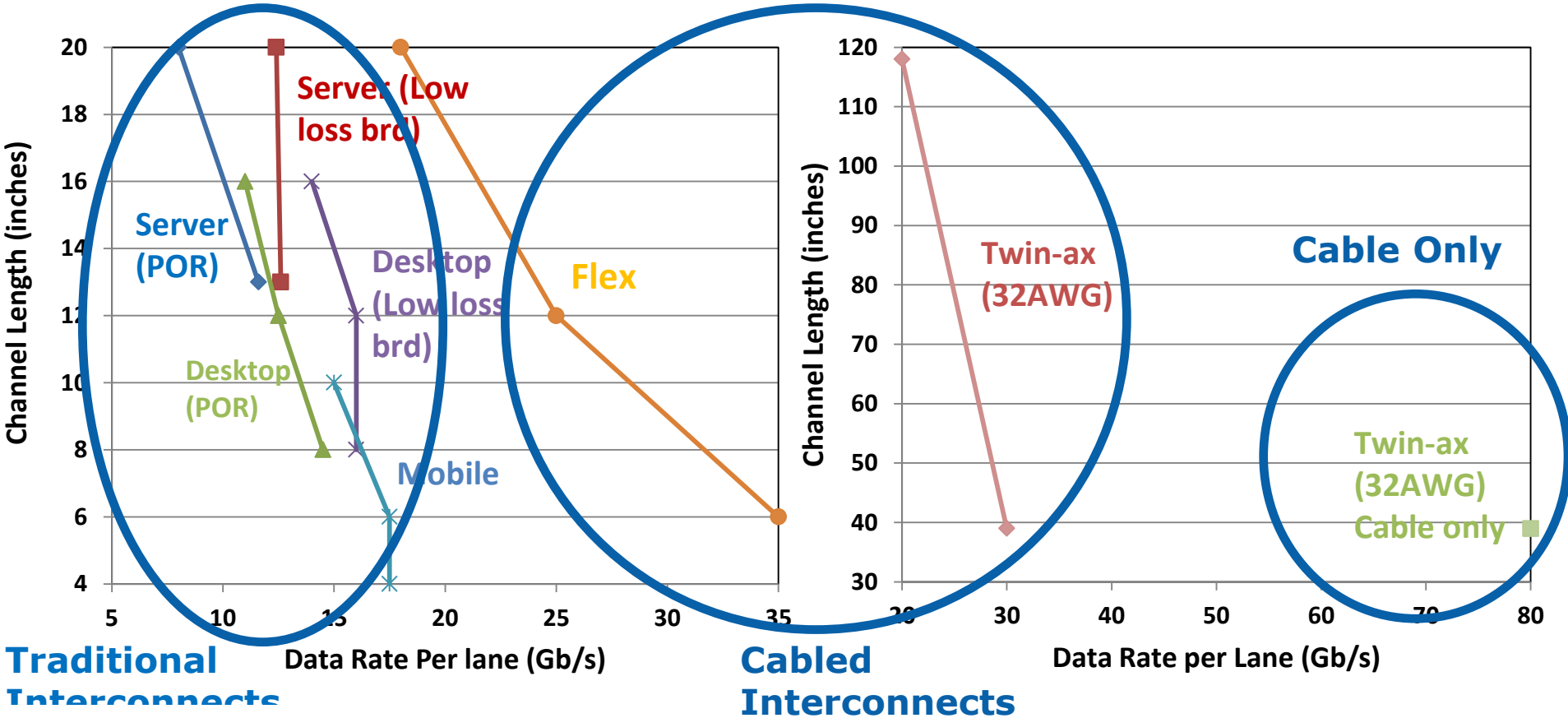
Flex



Twinax

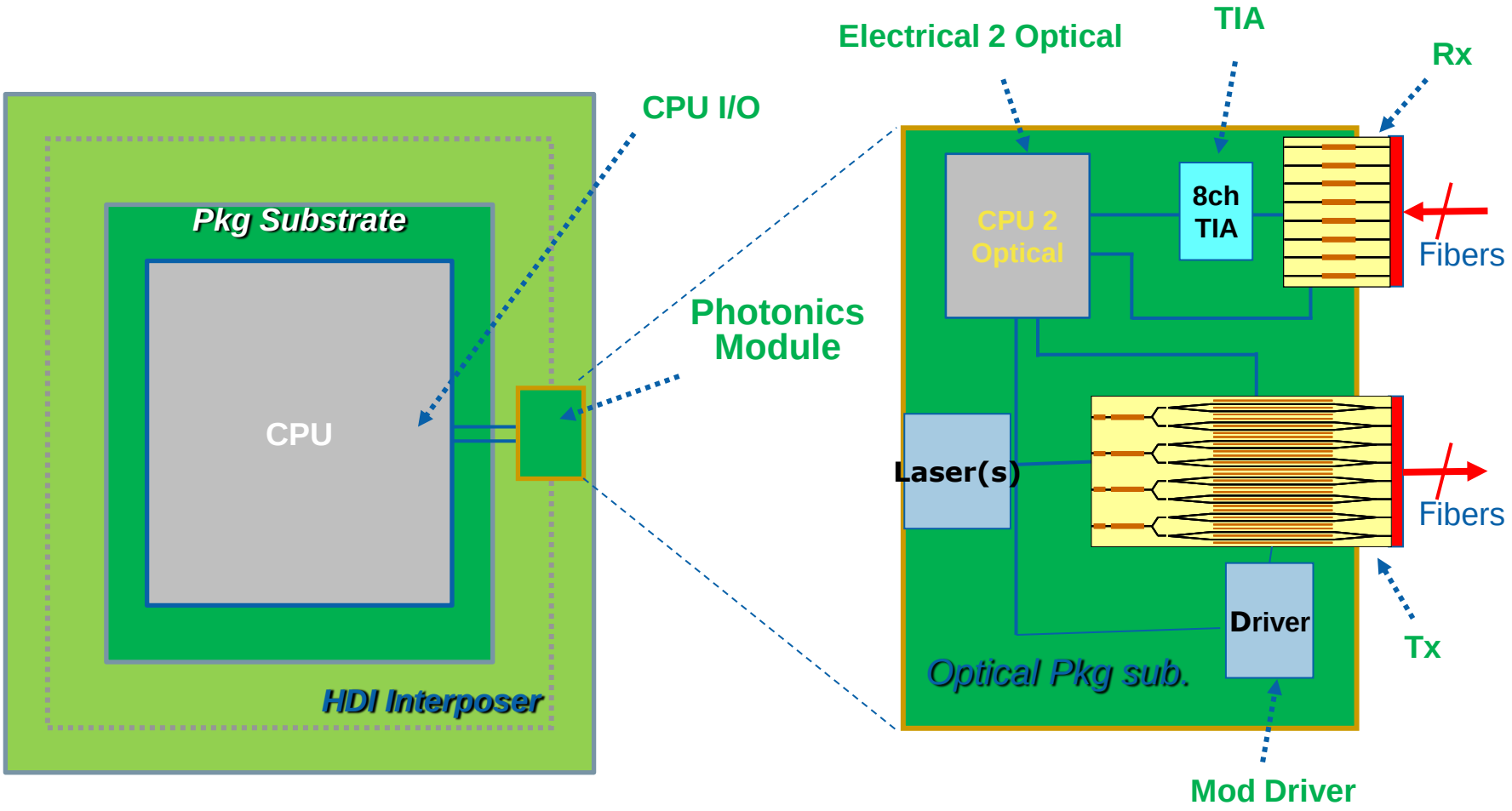


Channel Capacity vs. Distance



- Traditional interconnects limited to ~10-18Gb/s
- Top-of-package, cabled interconnects provide scaling to ~35Gb/s limited by packages, connectors etc.
- ***Research focused on achieving cable only capacity***

Photonics System



- ***Distance solution***
- ***Use when needed***

Summary

- Microprocessor I/O performance and power must scale
- Traditional interconnects nearing EOL
- 3D technology and dense interconnects compact the platform
- Short, dense electrical interconnect have high scalability
- Cabled electrical interconnect for medium distance also scale
- Electrical I/O research focused on realizing total available cable BW of 64Gb/s or greater
- Utilize active optical cables for distance > ~1m