Challenges and Solutions to DATA I/O

Keren Bergman

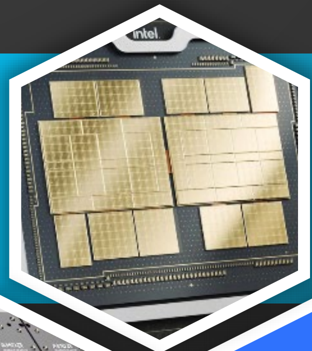
Professor of Electrical Engineering; Director, Columbia Nano Initiative
Columbia University

This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA).

The views, opinions and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

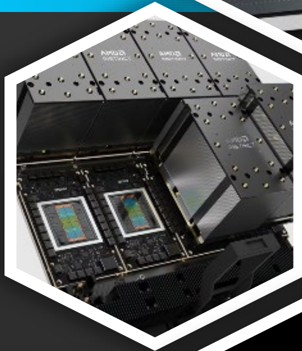
Distribution Statement A – Approved for public release. Distribution unlimited.

Challenges Moving Data Off-Chip



On-Chip

GPU-Memory Bandwidth
AMD MI300X: 5.2 TB/s
Nvidia DGX H100: 3.35 TB/s
Intel Data Center GPU Max: 3.28 TB/s

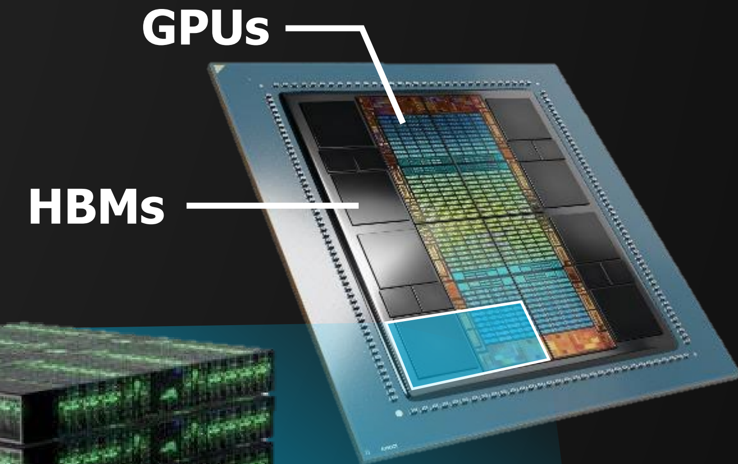


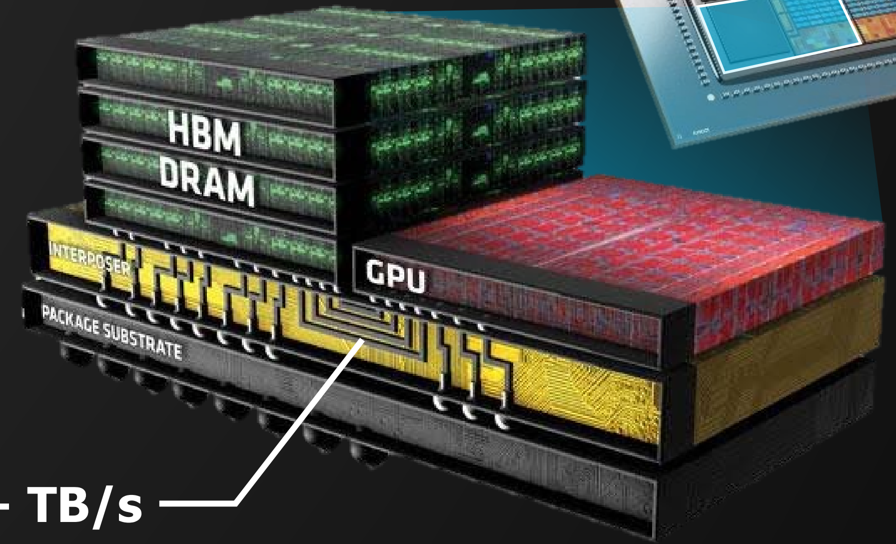
In-Socket

Off-Socket

GPUs

HBMs



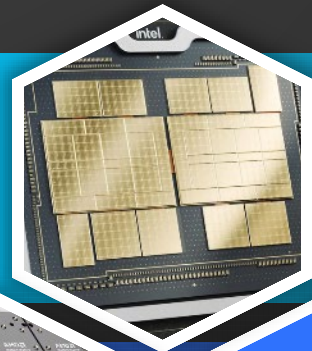


Electrical, up to 5+ TB/s

Images sources: AMD, Intel, and Nvidia.

Distribution Statement A – Approved for public release. Distribution unlimited.

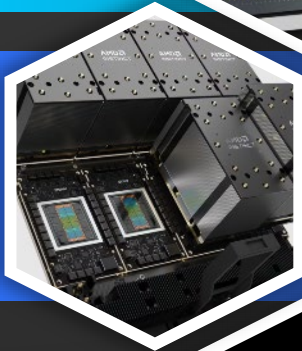
Challenges Moving Data Off-Chip



On-Chip

GPU-Memory Bandwidth

AMD MI300X: 5.2 TB/s
Nvidia DGX H100: 3.35 TB/s
Intel Data Center GPU Max: 3.28 TB/s

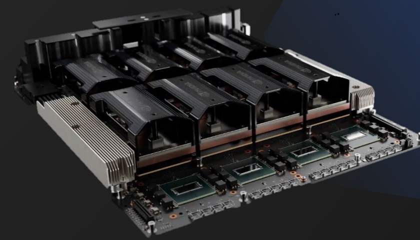


In-Socket

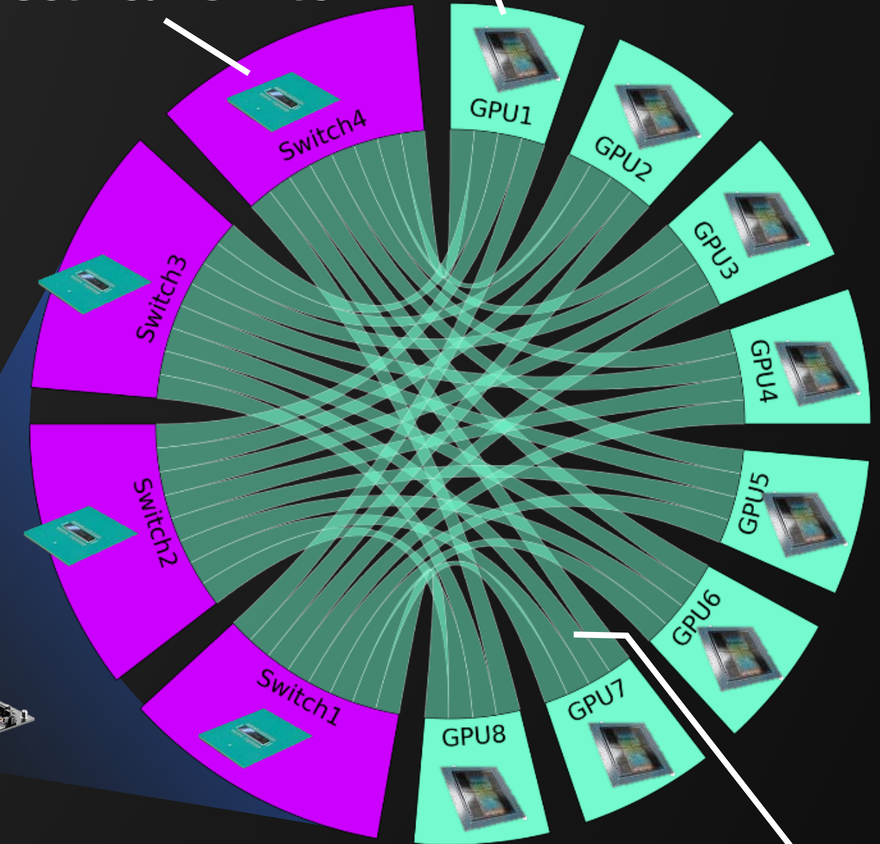
GPU-GPU Bandwidth

AMD Infinity Fabric: 800 GB/s
Nvidia NVLink & NVSwitch: 900 GB/s
Intel Xe Link: 720 GB/s

Off-Socket



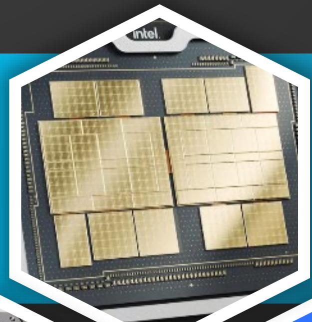
Multi-Chip Socket



Images sources: AMD, Intel, and Nvidia.

Distribution Statement A – Approved for public release. Distribution unlimited.

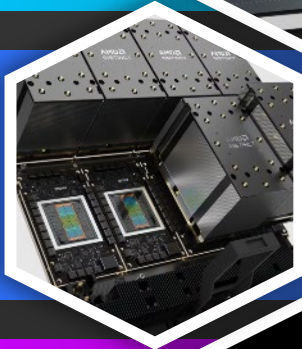
Challenges Moving Data Off-Chip



On-Chip

GPU-Memory Bandwidth

AMD MI300X: 5.2 TB/s
Nvidia DGX H100: 3.35 TB/s
Intel Data Center GPU Max: 3.28 TB/s



In-Socket

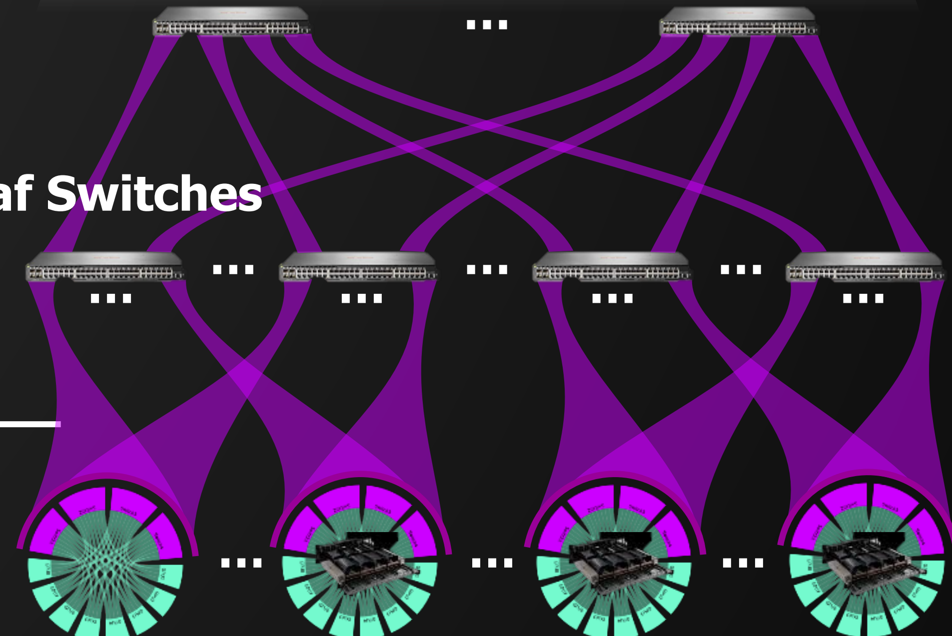
GPU-GPU Bandwidth

AMD Infinity Fabric: 800 GB/s
Nvidia NVLink & NVSwitch: 900 GB/s
Intel Xe Link: 720 GB/s

Off-Socket

Off-Socket Link Bandwidth

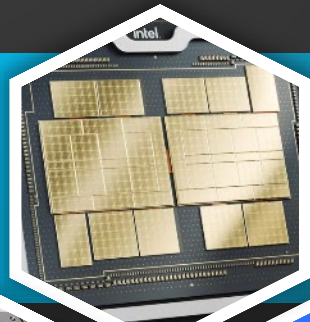
InfiniBand: 400 Gb/s
Projected 800 Gb/s near future



Images sources: AMD, Intel, and Nvidia.

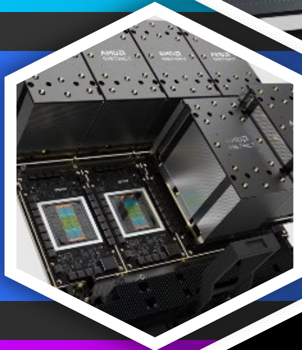
Distribution Statement A – Approved for public release. Distribution unlimited.

Challenges Moving Data Off-Chip



On-Chip

GPU-Memory Bandwidth
AMD MI300X: 5.2 TB/s
Nvidia DGX H100: 3.35 TB/s
Intel Data Center GPU Max: 3.28 TB/s



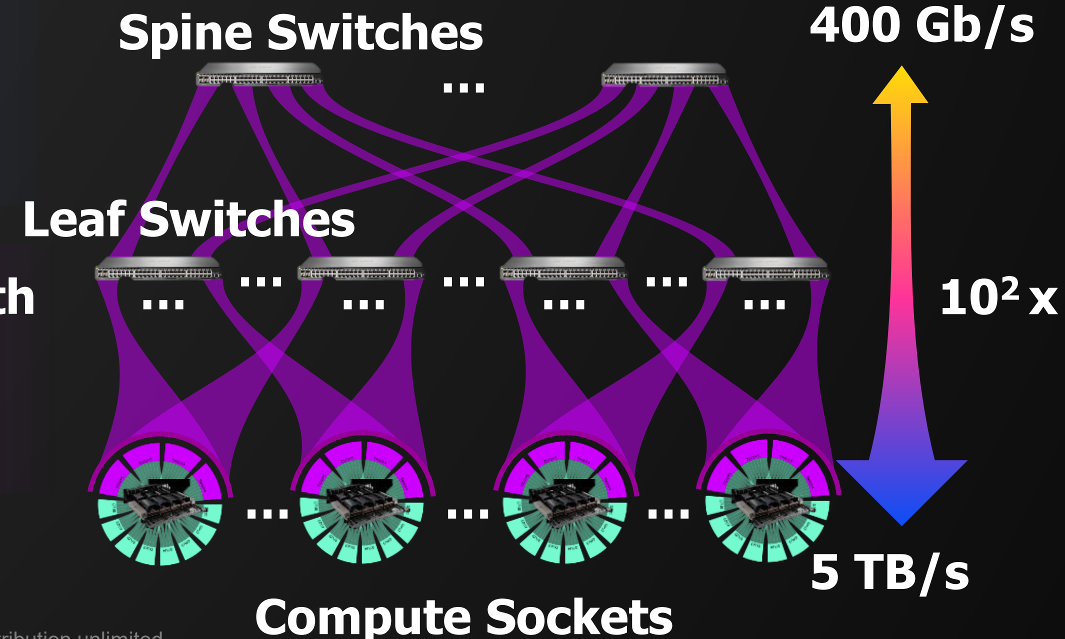
In-Socket

GPU-GPU Bandwidth
AMD Infinity Fabric: 800 GB/s
Nvidia NVLink & NVSwitch: 900 GB/s
Intel Xe Link: 720 GB/s

Off-Socket

Off-Socket Link Bandwidth
InfiniBand: 400 Gb/s
Projected 800 Gb/s near future

Off-Socket IO BW limit creates **100 X Bandwidth Taper** across system

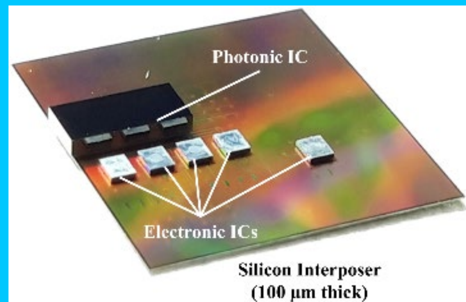


Images sources: AMD, Intel, and Nvidia.

Distribution Statement A – Approved for public release. Distribution unlimited.

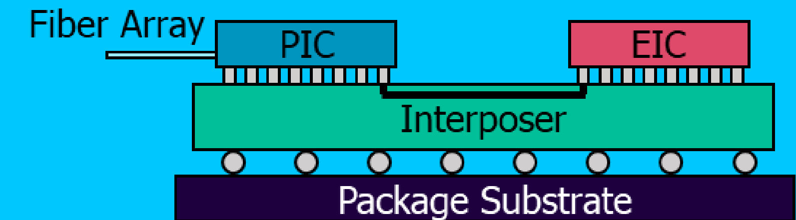
Bringing Photonics to the Chip

2.5D Integration



Pros:
- Better density than 2D
- Balanced scalability & flexibility
- Thermal isolation

Cons:
- Parasitics from doubled bump interfaces and traces
- Still limited BW density
- Added complexity from interposer design



2.5D Integration

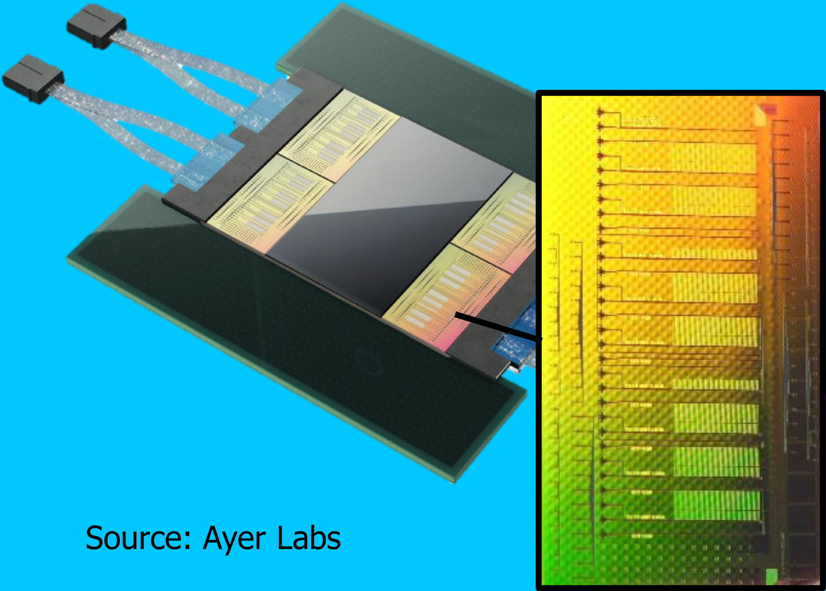
~400 Gbps/mm

~10 pJ/b

Distribution Statement A – Approved for public release. Distribution unlimited.

Bringing Photonics to the Chip

Monolithic Integration



Source: Ayer Labs

Pros:

- Minimal parasitics
- Simplified packaging
- Thermal dissipation

Cons:

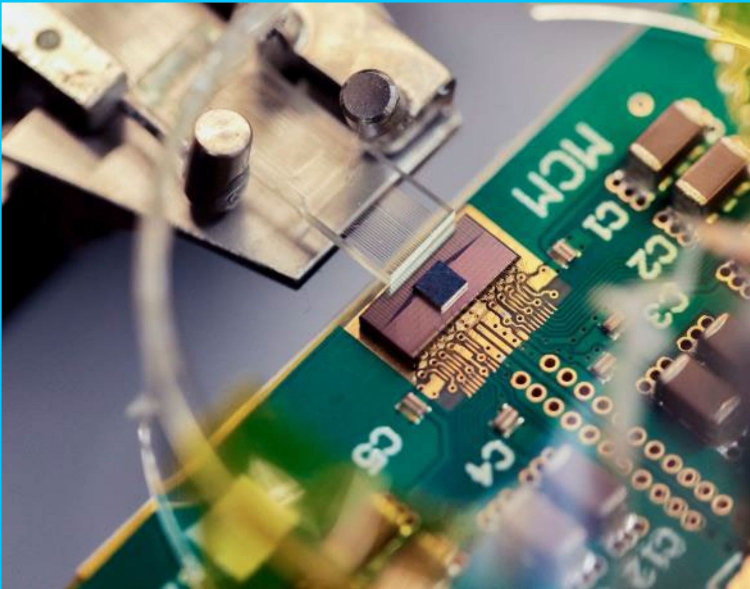
- Bandwidth density limited by electronics
- Outdated technology nodes limit power, scaling

2.5D Integration	Monolithic Integration	
~400 Gbps/mm ~10 pJ/b	~200 Gbps/mm ~5 pJ/b	

Distribution Statement A – Approved for public release. Distribution unlimited.

Bringing Photonics to the Chip

3D Integration

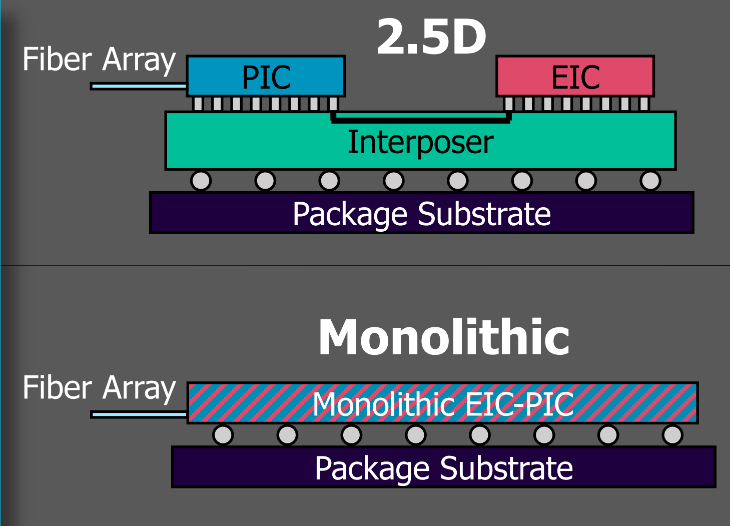


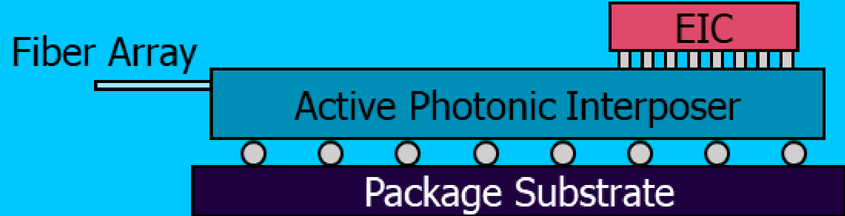
Advantages:

- Best shoreline & area bandwidth density
- Massive wavelength scalability
- Benefits from advanced CMOS technology nodes

Challenges:

- Packaging yield
- Thermal management

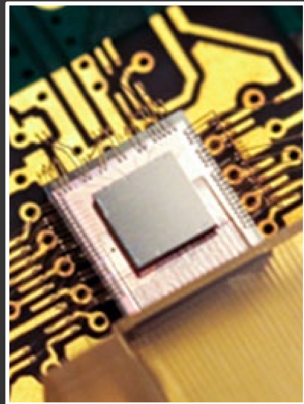


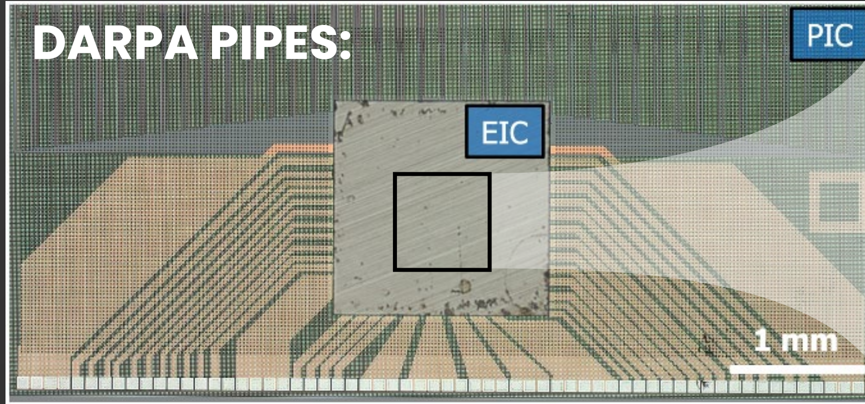


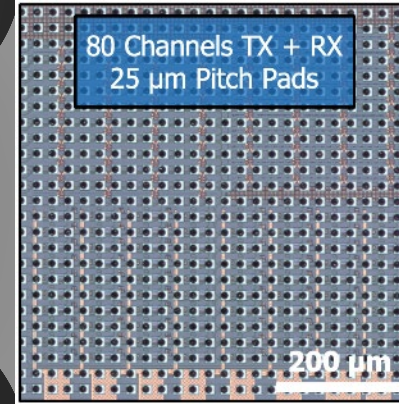
2.5D Integration	Monolithic Integration	3D Integration
~400 Gbps/mm	~200 Gbps/mm	Multi-Tbps/mm
~10 pJ/b	~5 pJ/b	Sub-pJ/b

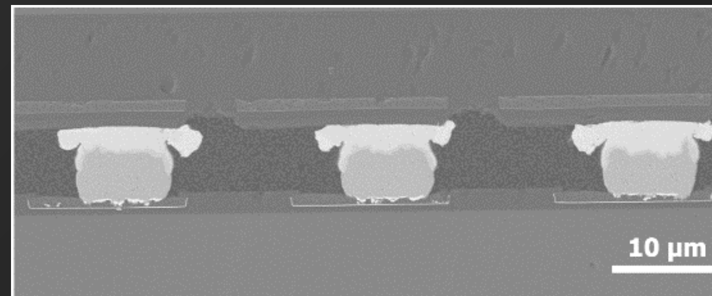
Distribution Statement A – Approved for public release. Distribution unlimited.

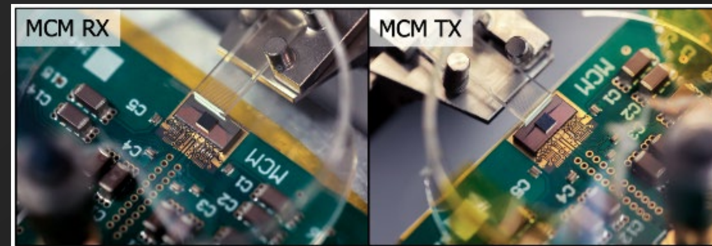
3D Integration for Scalable Optical I/O





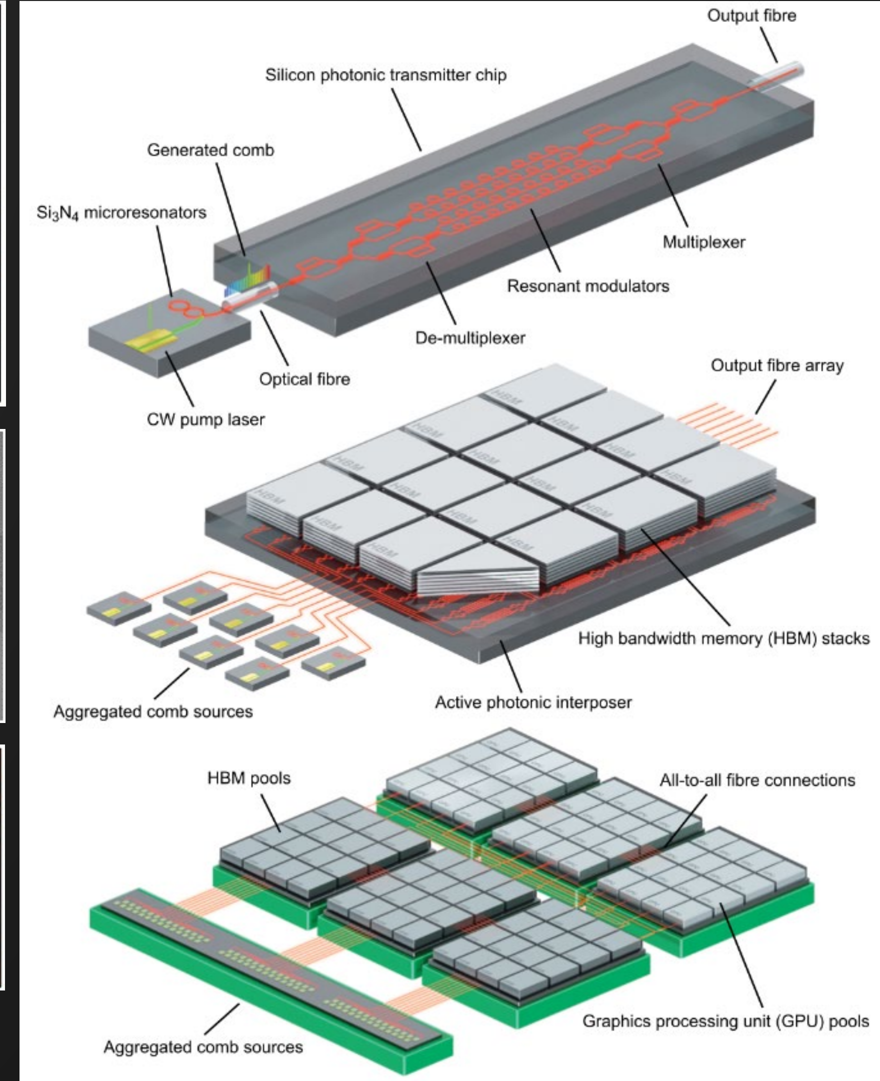






5.3 Tbps/mm²
@ 120 fJ/b

Source: S. Daudlin *et al.*, in OFC 2023, M3I.1



Source: A. Rizzo *et al.*, Nature Photonics, 2023

Distribution Statement A – Approved for public release. Distribution unlimited.

THANK YOU