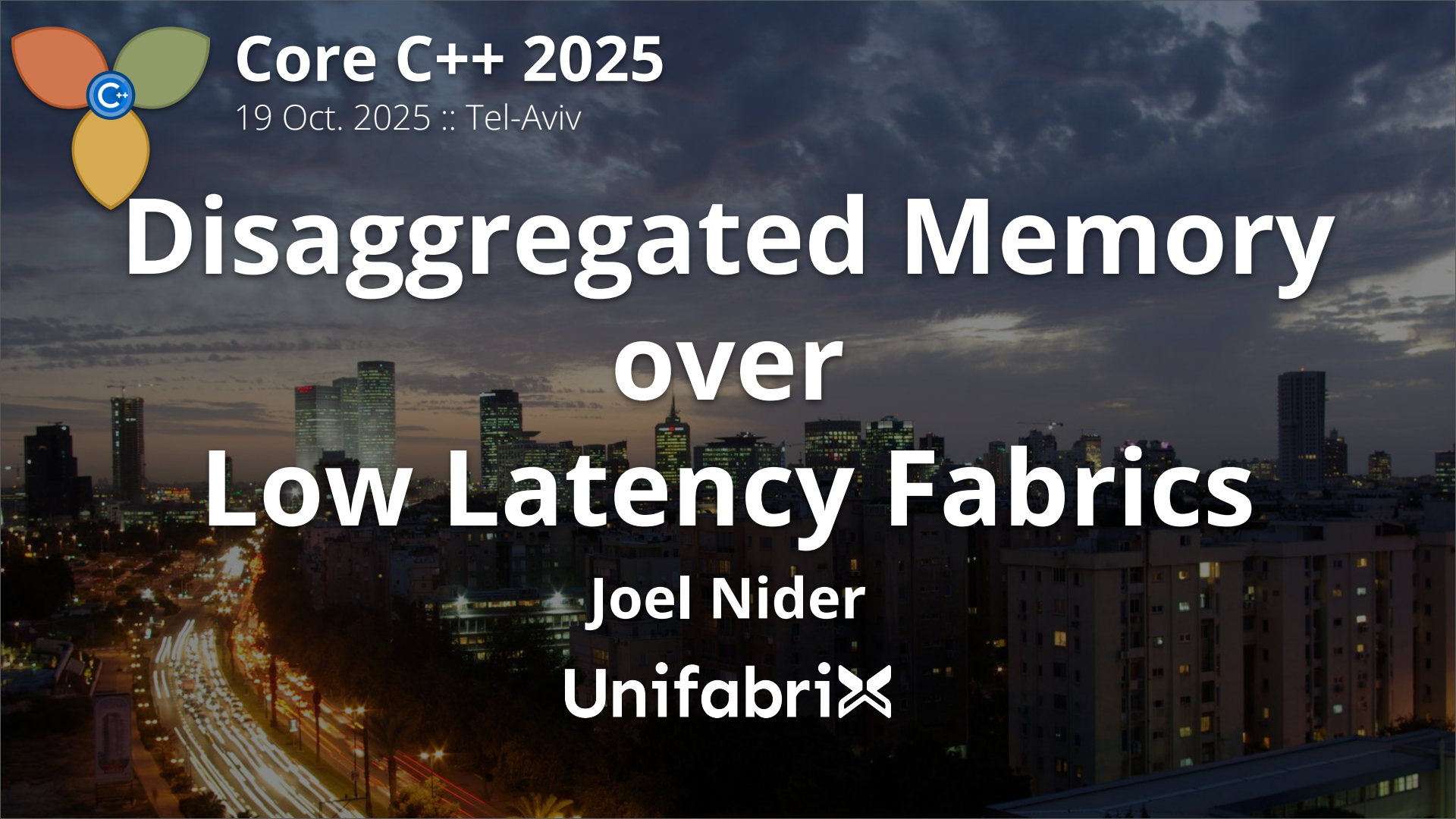

Core C++ 2025
19 Oct. 2025 :: Tel-Aviv
Disaggregated Memory over Low Latency Fabrics
Joel Nider
Unifabrix

Big problems require big solutions

High Performance Computing (Supercomputers)

AI Machine Learning (Custom clusters)

Hyperscaler Cloud Computing (SaaS)

When one computer won't do

Two computers are better than one

solution: connect multiple computers with a network

We need a better interconnect

Networks are not designed for memory access

- We have known this for a long time
- Explicit communication
  - Software APIs (e.g. sockets, RDMA)
- Emphasis on bandwidth
  - large transfers are more efficient

Memory Fabrics

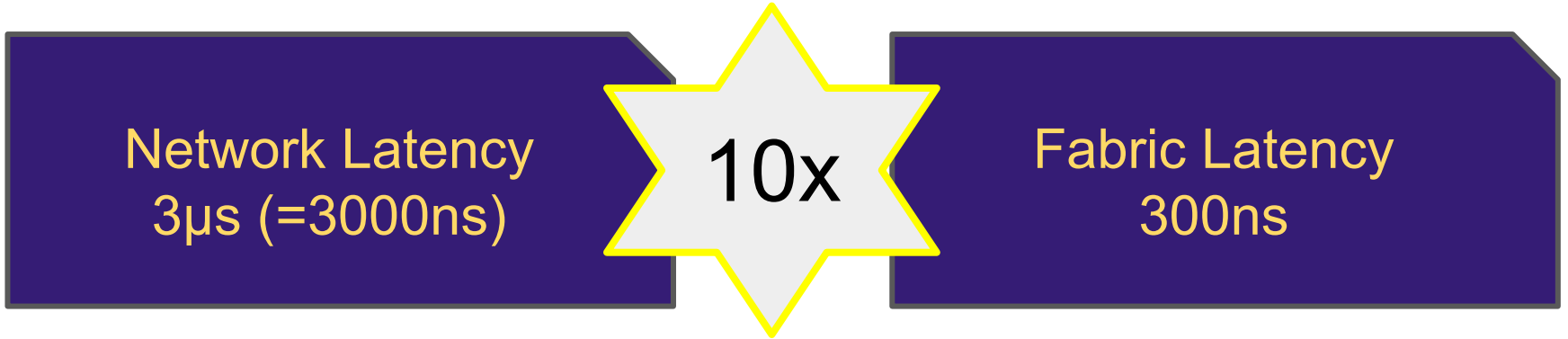
Memory Semantics

Load + Store Instructions
Ordering

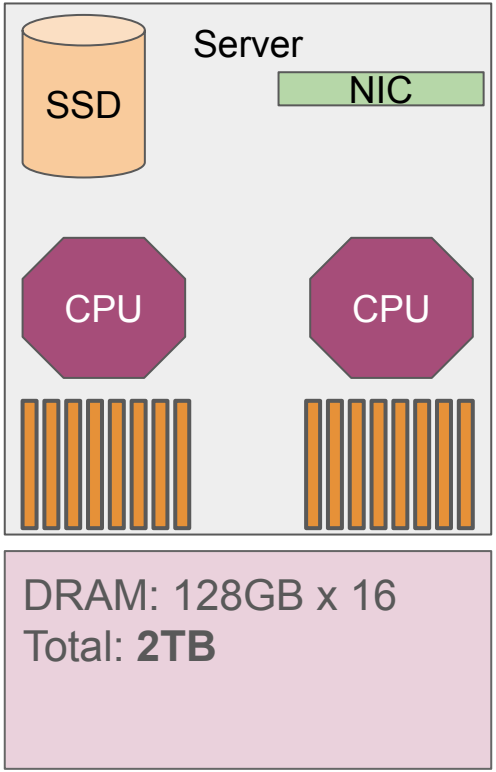
Cache Coherency

[Nice to have, not Mandatory]

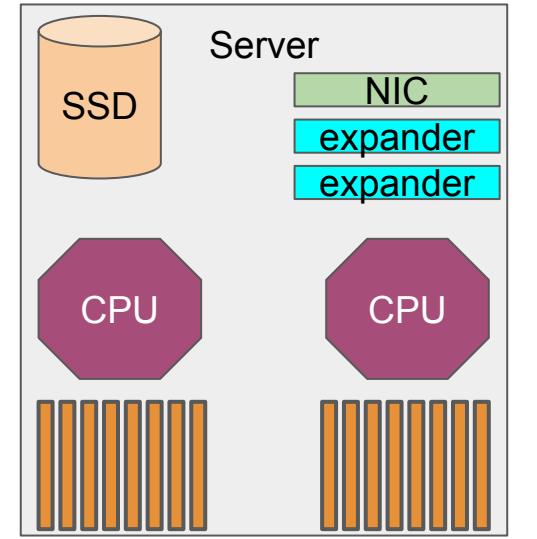
Fundamentally different software from networks



Fabric-Attached Memory

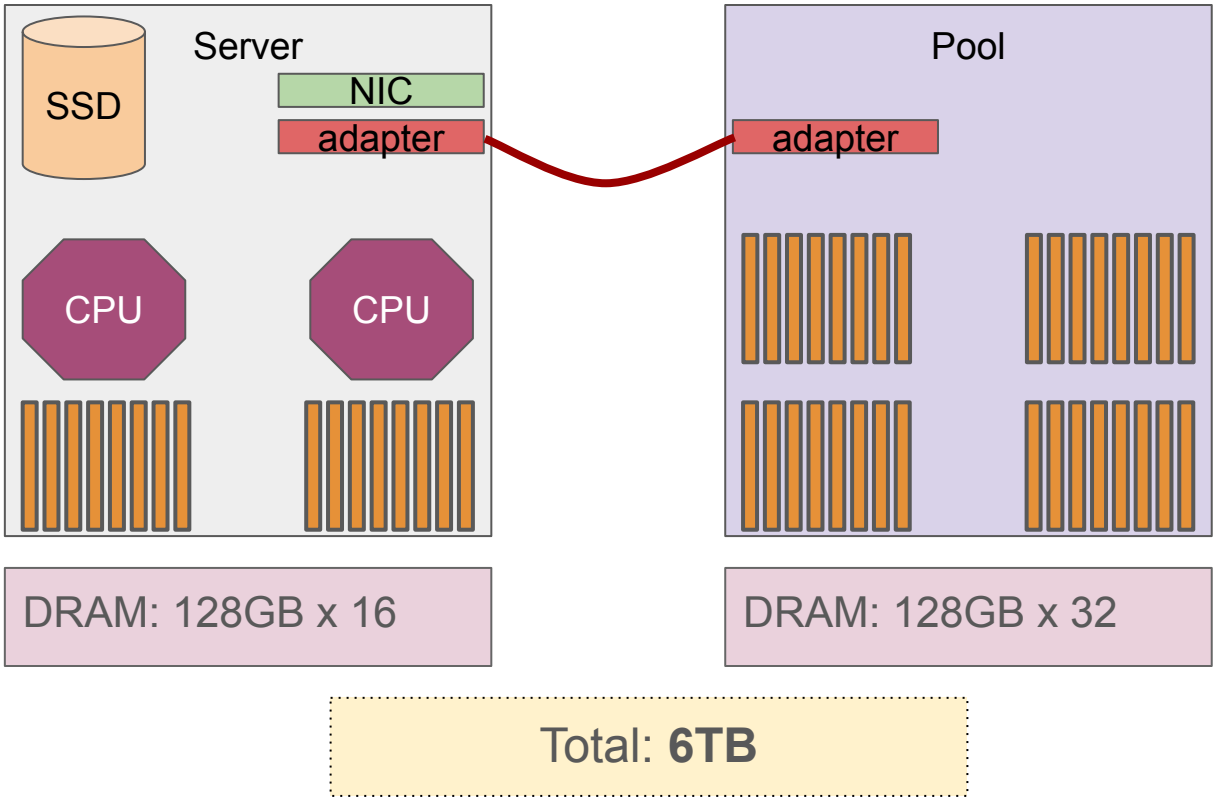


Expander

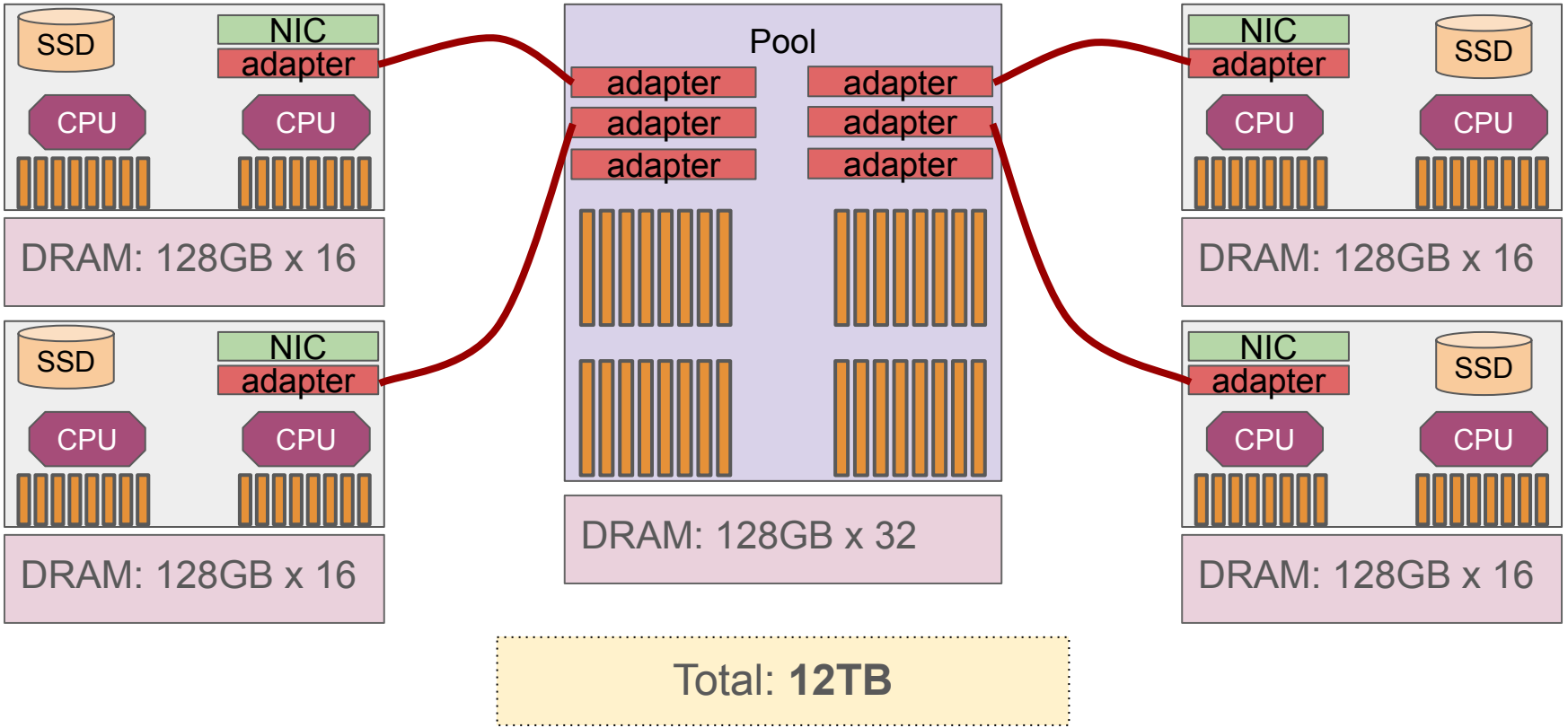


DRAM: 128GB x 16
Expander: 512GB x 2
Total: **3TB**

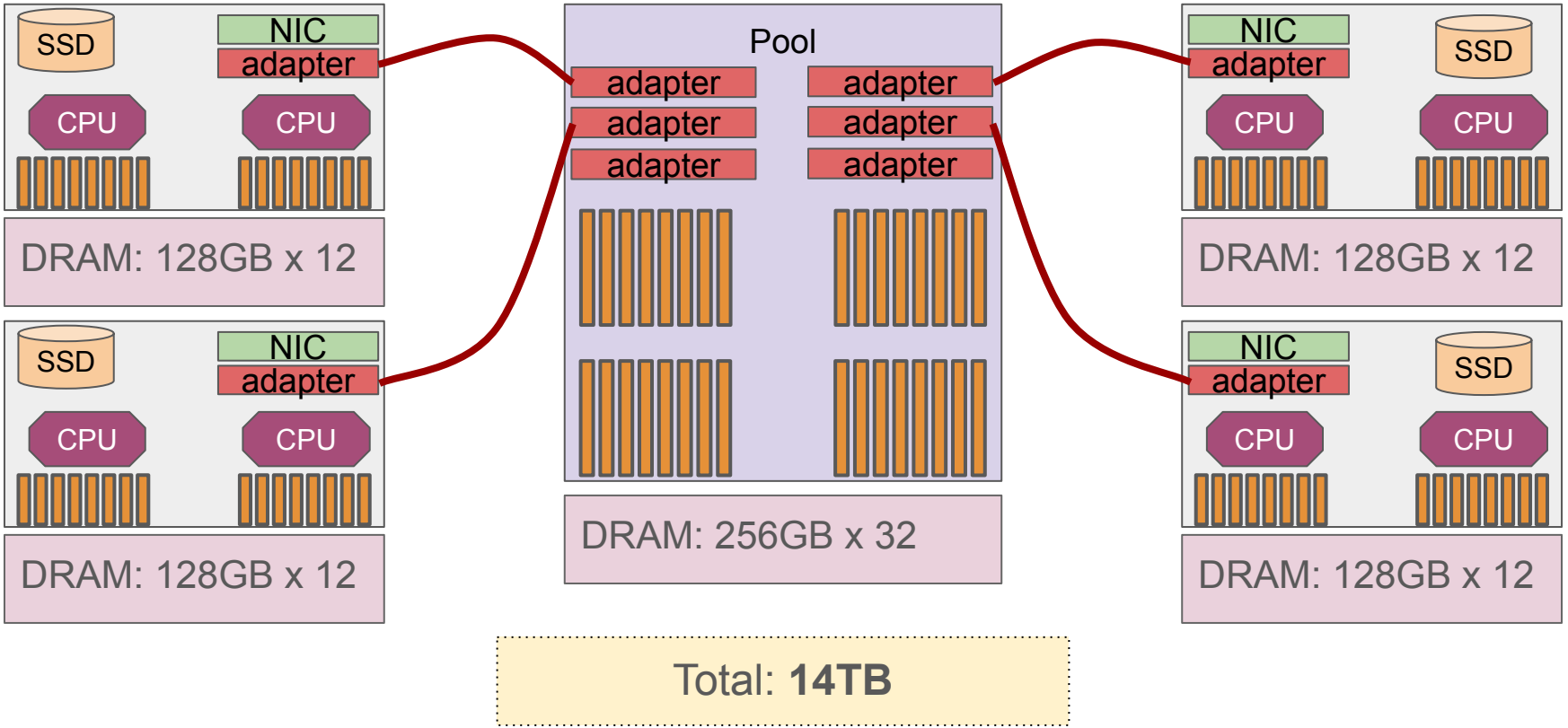
Pool


Server
SSD
NIC
adapter
CPU
CPU
Pool
adapter
DRAM: 128GB x 16
DRAM: 128GB x 32
Total: 6TB

Pool


SSD
NIC
adapter
CPU
CPU
DRAM: 128GB x 16
SSD
NIC
adapter
CPU
CPU
DRAM: 128GB x 16
Pool
adapter
adapter
adapter
adapter
adapter
adapter
DRAM: 128GB x 32
NIC
adapter
SSD
CPU
CPU
DRAM: 128GB x 16
NIC
adapter
SSD
CPU
CPU
DRAM: 128GB x 16
Total: 12TB

Shared Memory Pool



Memory Fabrics

Many have tried, none have succeeded

CXL

CXL [November 2020]

Members: **80**

Physical Layer: **PCIe 5**

Coherency: **yes**

The Battle Of the Protocols

- CXL - announced 2020
- UALink - announced in April 2025
- NVLink Fusion - announced in May 2025
- Scale-Up Ethernet - announced in May 2025

Who will win?

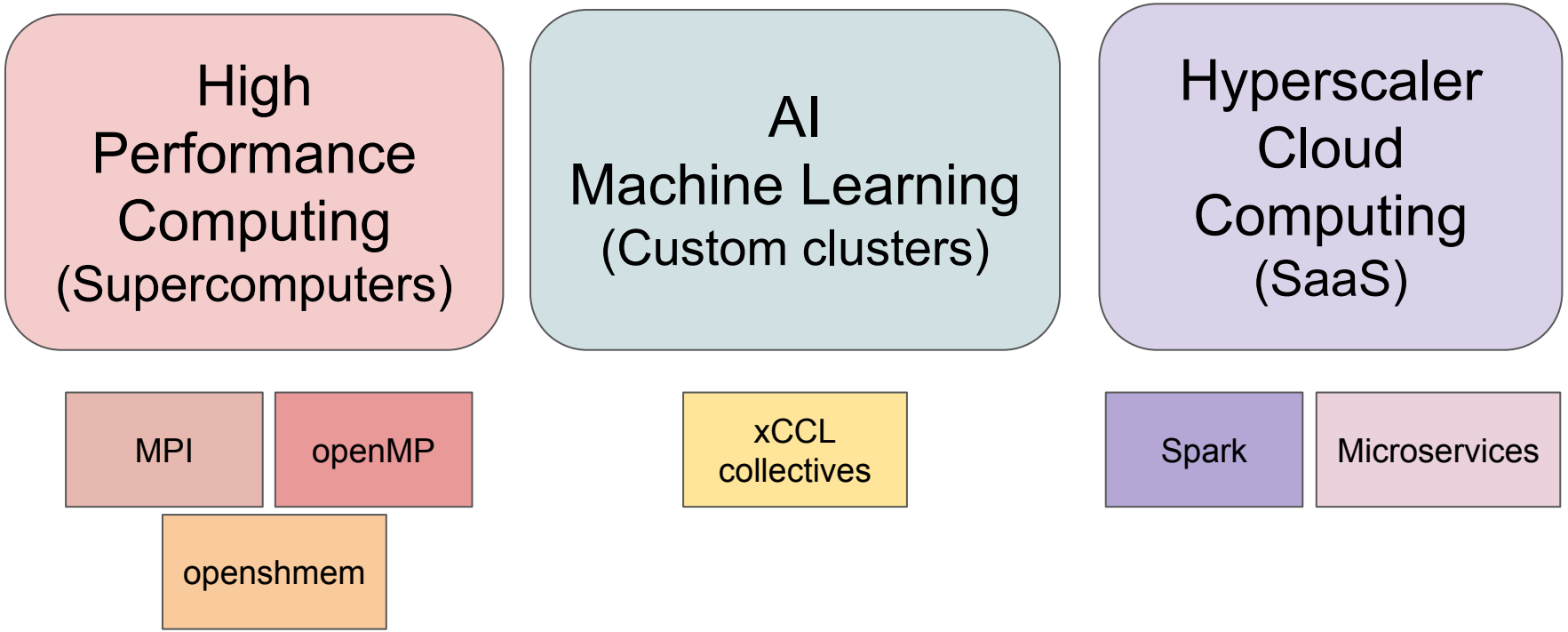
UALink [April 2025]

Members: **69**

Physical Layer: **XXXX**

Coherency: **no**

Use Cases


High Performance Computing (Supercomputers)
AI Machine Learning (Custom clusters)
Hyperscaler Cloud Computing (SaaS)
MPI
openMP
openshmem
xCCL collectives
Spark
Microservices

Software Solutions

API for controlling the remote memory

- Allocate/free
- Access control
- Notifications

Necessary primitives

- Atomics
- Barriers

Summary

- Disaggregated memory is now a reality
- Memory expanders, pools, and shared pools are changing system design in fundamental ways
- CXL is successful but has competition
- The winner(s) have yet to be decided

More memory is better

Unifabrix

Cache Coherency

- Cache is used to hold a local copy of data
  - Generally lower latency than main memory
  - Shorter distance for data movement
- Cached shared memory requires cache coherency
  - If two compute units update their local copies simultaneously, who wins?
- Some protocols do not support coherency by design (NVLink, UALink)
- Hardware coherency has not yet been implemented by chip vendors (CXL)
- Rely on software coherency for now

NVLink Fusion [May 2025]

Members: **NVidia**

Physical Layer: **XXXX**

Coherency: **yes**

Scale-Up Ethernet [May 2025]

Members: **Broadcom**

Physical Layer: **Ethernet**

Coherency: **yes**