CREDO
we connect.

Credo OmniConnect (Weaver)

Haoli Qian

Chief R&D Officer

CONFIDENTIAL

Credo FY2027 Q1 at a Glance
LTM

$1.6B
LTM Revenue

115%
Q1 YoY Revenue Growth

24+
Tier 1 Customers

68%
FY27Q1 Non-GAAP Gross Margin[1]

Marquee Engagements
Top Hyperscale Data Center Operators

48.2%
FY27Q1 Non-GAAP Operating Margin[1]

969
Global Personnel (as of Jul'26)



*

FY25 Rev: $437M
FY26 Rev: $1.3B
FY27 Rev: ~$2.5B
(Analyst consensus)

Note: Fiscal year ends the Saturday closest to each April 30th. Product revenue represents product sales and product engineering services. 1 Reflects non-GAAP financial measure For a reconciliation of this non-GAAP financial measure to the most comparable GAAP measure, please see the appendix.

CONFIDENTIAL

Global Operations Support Leading Development

World-class DSP and analog/mixed capability for best performance, power, and cost



985

Employees across USA, Canada, Taiwan, China, Singapore, Israel

Worldwide

R&D, Design, Sales & Support reps

TAIWAN FABRICATION PARTNER

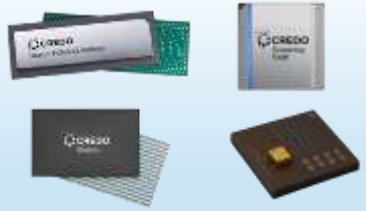
CONFIDENTIAL

CREDO 2.0
Connectivity Solutions

Active Electrical Cables
Up to 7m



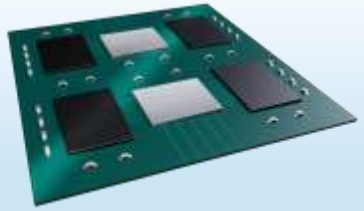
PCIe, Ethernet Retimers
& Optical DSPs



ZeroFlap Transceivers
Up to 2km



OmniConnect
Gearboxes
Up to 250mm



Active LED Cables
Up to 30m



PILOT Debug &
Telemetry Software



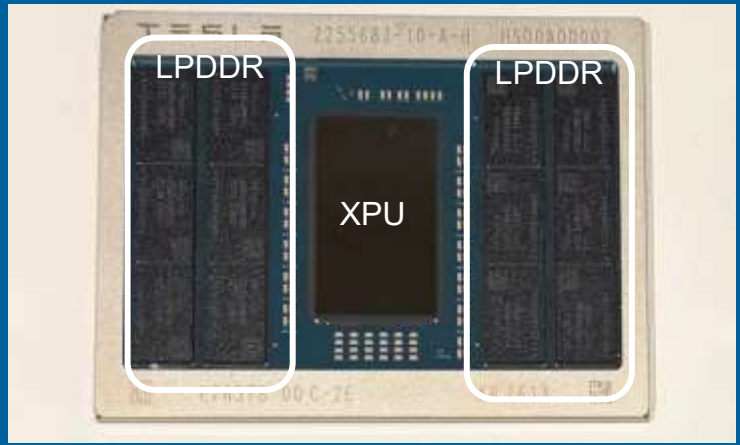
CONFIDENTIAL

Issue

The Inference Memory Wall

- In AI inference workloads:
  - Throughput is memory bound, not FLOP bound
- Limitation w/ existing solutions (LPDDRx, GDDRx, HBM):
  - Beachfront Barrier
    - Consume excess beachfront
  - Substrate Barrier
    - Too little memory on substrate

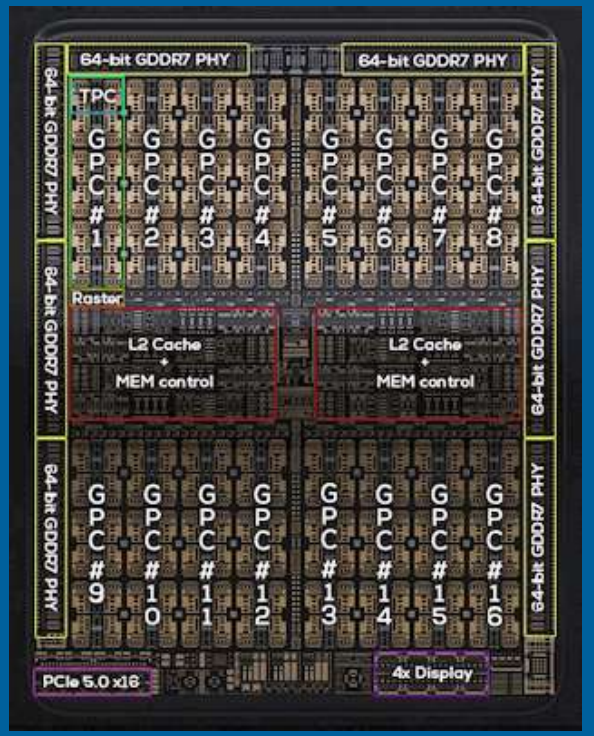
Substrate Barrier



Organic Substrate 100x100

HBM4 = 144 GBytes / RSD
LPDDR5X = 512 GBytes

Beachfront Barrier



LPDDR5X = ~0.18 Tbps/mm
GDDR7 = ~0.4 Tbps/mm
HBM4 = ~2 Tbps/mm

CONFIDENTIAL

LPDDR vs HBM comparison

Metric	LPDDR	HBM
Memory Bandwidth	**500GB/s~1TB/s for max die**	**5~10TB/s for single GPU die**
Reach distance	**20~30mm PCB**	**3~5mm on silicon interposer**
Packaging	Standard or advanced packaging	Must use advanced packaging
Thermal management	Easy and low risk allows air cooling	Very challenging with 8/12/16 stacked dies need liquid cooling
DRAM foundry manufacturing capacity consumption	1x/GB	3~4X/GB (even for mature HBM manufacturing)
Cost	1x	5~6x
Memory size per AI inference engine potential	2TB	200GB
AI system complexity due to cache size constraint	low	high
AI system total cost of ownership	low	high

CONFIDENTIAL

AI Applications Depend on Product of IO Density and IO Reach

Memory Bandwidth per Die

↓

IO Beachfront Density

×

Memory Connectivity Reach

=

Fanout

Drives AI system scale and cost

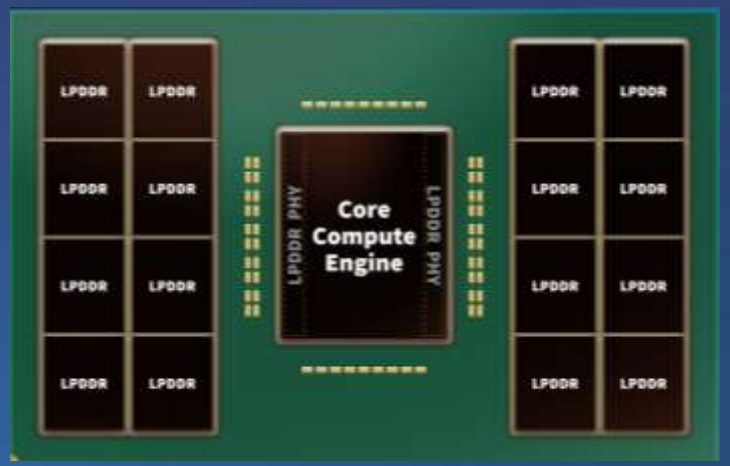


CONFIDENTIAL

OmniConnect: Best Memory Fanout Solution

Engineered to address AI Inference Memory Wall

- OmniConnect key IP: High beachfront density X 200-300mm reach VSR SerDes
- System solution:
  - Key VSR SerDes IP offered at various foundry process nodes to AI ASIC design partners
  - Weaver chiplet with 112G VSR SerDes to standard LPDDR5X
- Up to 12.8TB memory density and 8TB/s bandwidth with single XPU for AI inference



Integrated LPDDR



LPDDR Disaggregation w/ Weaver

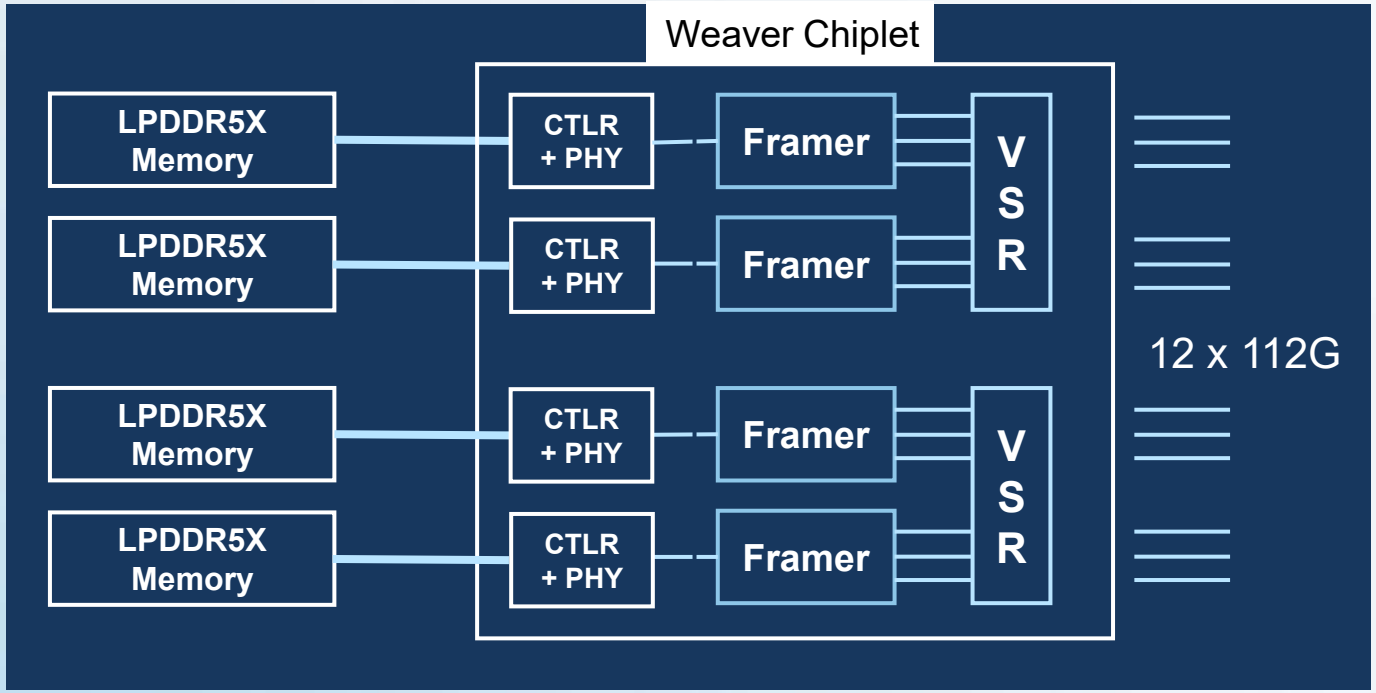
CONFIDENTIAL

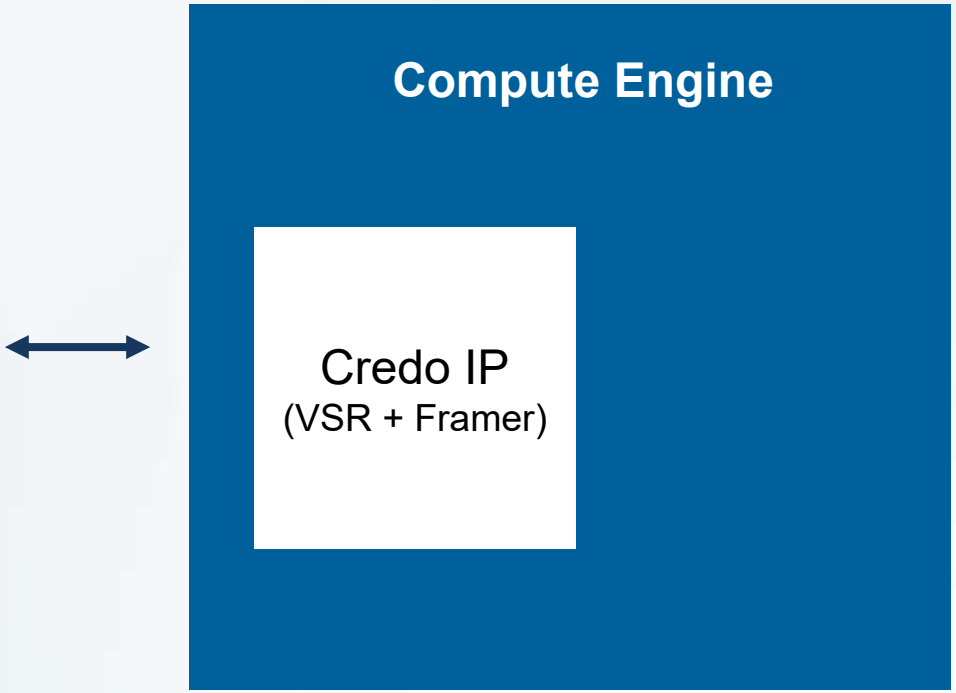
LPDDR vs HBM vs OmniConnect comparison

Metric	LPDDR	HBM	OmniConnect
Memory Bandwidth	500GB/s~1TB/s for max die	5~10TB/s for single GPU die	5~10TB/s for single GPU die
Reach distance	20~30mm PCB	3~5mm on silicon interposer	200-300mm PCB
Packaging	Standard or advanced packaging	Must use advanced packaging	Standard or advanced packaging
Thermal management	Easy and low risk allows air cooling	Very challenging liquid cooling	Easy and low risk allows air cooling
DRAM foundry manufacturing capacity consumption	1x/GB	3~4X/GB	1x/GB
Cost	1x	5~6x	1x
Memory size per AI inference engine potential	2TB	200GB	10TB
AI system complexity due to cache size constraint	low	high	low
AI system total cost of ownership	low	high	low

CONFIDENTIAL

OmniConnect with Weaver Block Diagram





Weaver Chiplet
- 12 VSR lanes
- 4 x32 LPDDR5x memory

Each IP macro → 6 VSR lanes

Process Nodes
- TSMC N3, N5, N7, N12

CONFIDENTIAL

Weaver

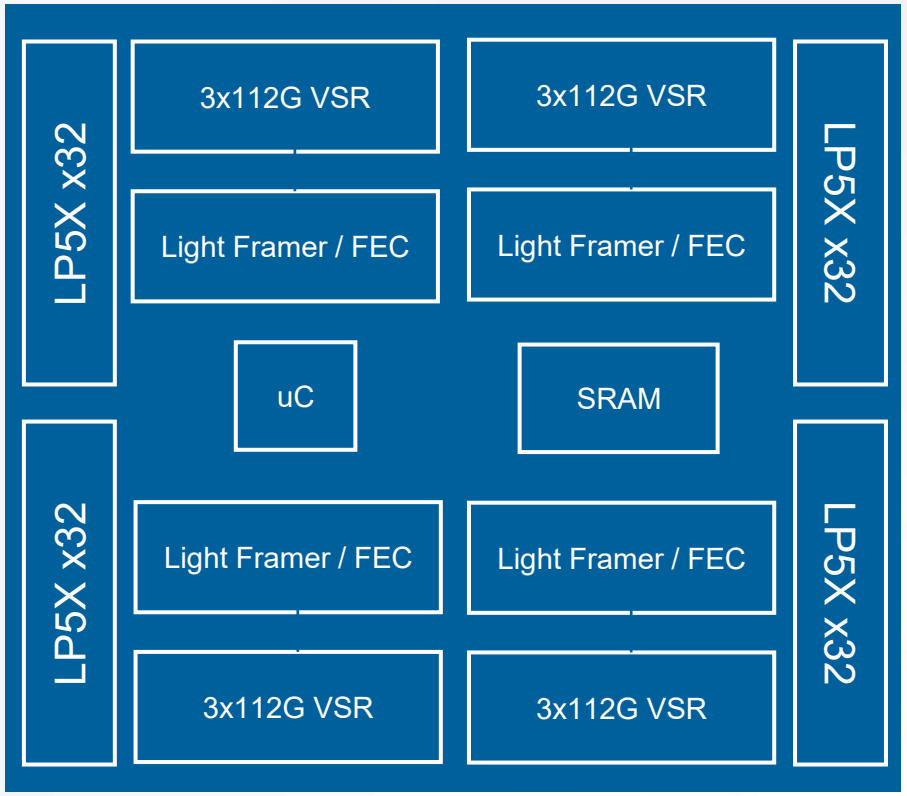
Credo solution

- 12x112G VSR (23dB IL @ 26GHz) to host
- Each die includes four macros – static lane – DRAM mapping
- Lightweight Framer / FEC follows PCIe6 FLIT
  - Credit based flow control
  - FEC / CRC from PCIe6 FLIT
  - 70ns RTT latency including FEC
  - Full control over LPDDR5X PHYs (page status / refresh / etc.)
- Bandwidth per Weaver
  - 168 GB/s raw BW – Over VSR Serdes
  - 155 GB/s memory throughput
- N5 process node

Schedule

- Taped out July/26
- Samples Oct/26

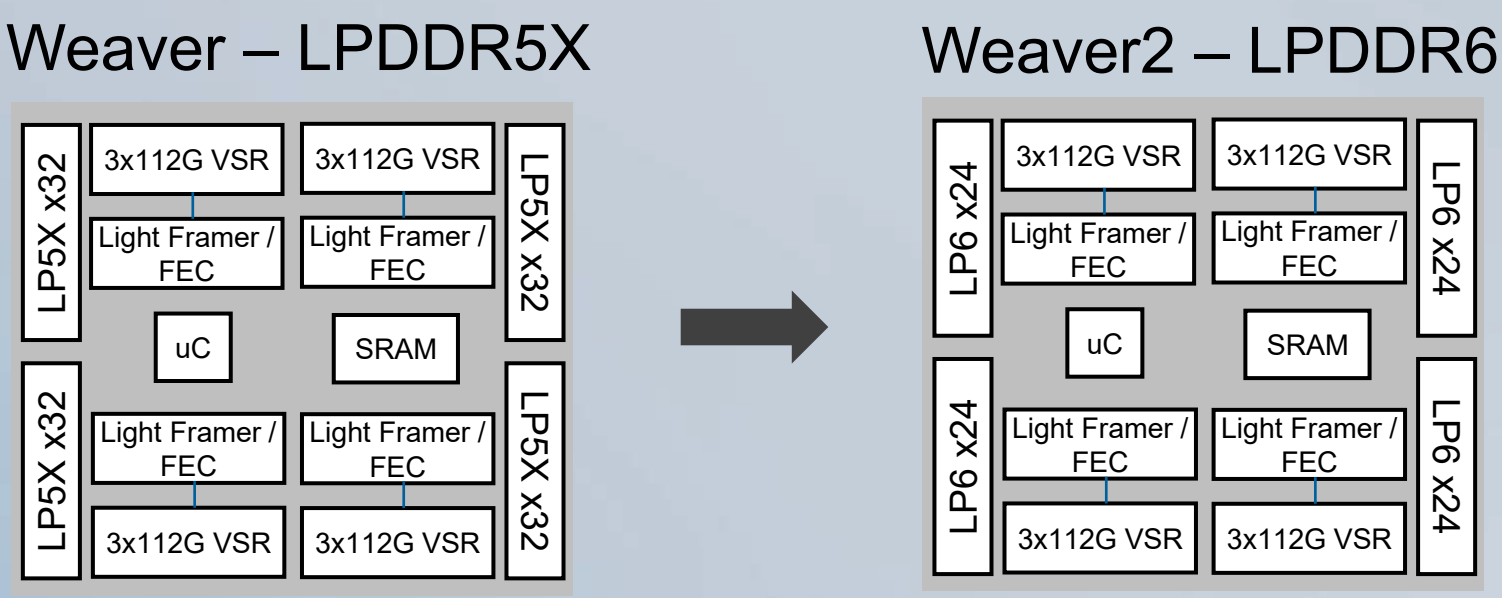
Weaver N5 Floorplan



CONFIDENTIAL

LPDDR5X to LPDDR6 Transition

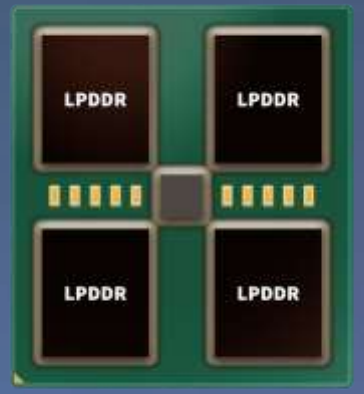
No Change in Host



- LPDDR5X x32 @ 9.6Gb/s = LPDDR6 x24 @ 14.4Gb/s – simple HW migration
- No change required to framer or host hardware interface

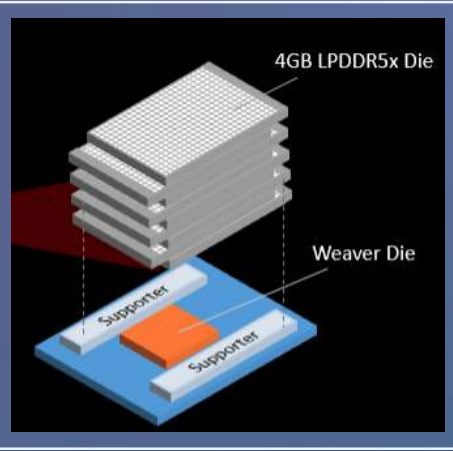
- Credo will offer Weaver2 in mid 2027 once LPDDR6 ecosystem matures

External Memory
Packaging Options



Weaver + DRAM on substrate

~33x35mm with 4 x32 LP5X



High Density Compact Package with Integrated Weaver

High density in small package footprint



Packaged Weaver



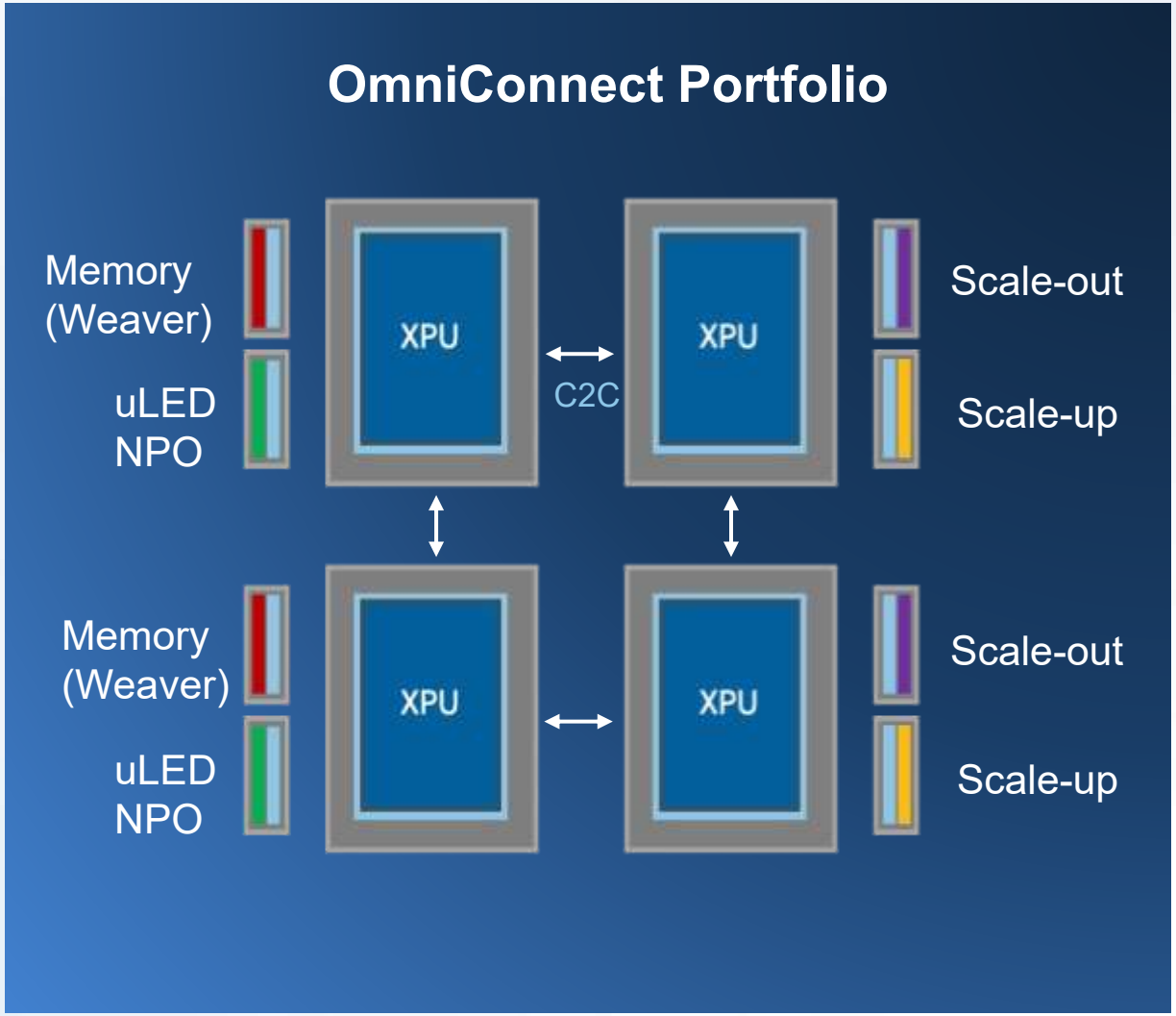
DIMM Connector

CONFIDENTIAL

Credo OmniConnect – ScaleUp/ScaleOut/Memory

OmniConnect

- A general-purpose AXI-over-VSR bus
- Can be late-bound to multiple peripherals without needing CoWoS



CONFIDENTIAL

CREDO
we connect.

Thank you

CONFIDENTIAL