



EVIDEN



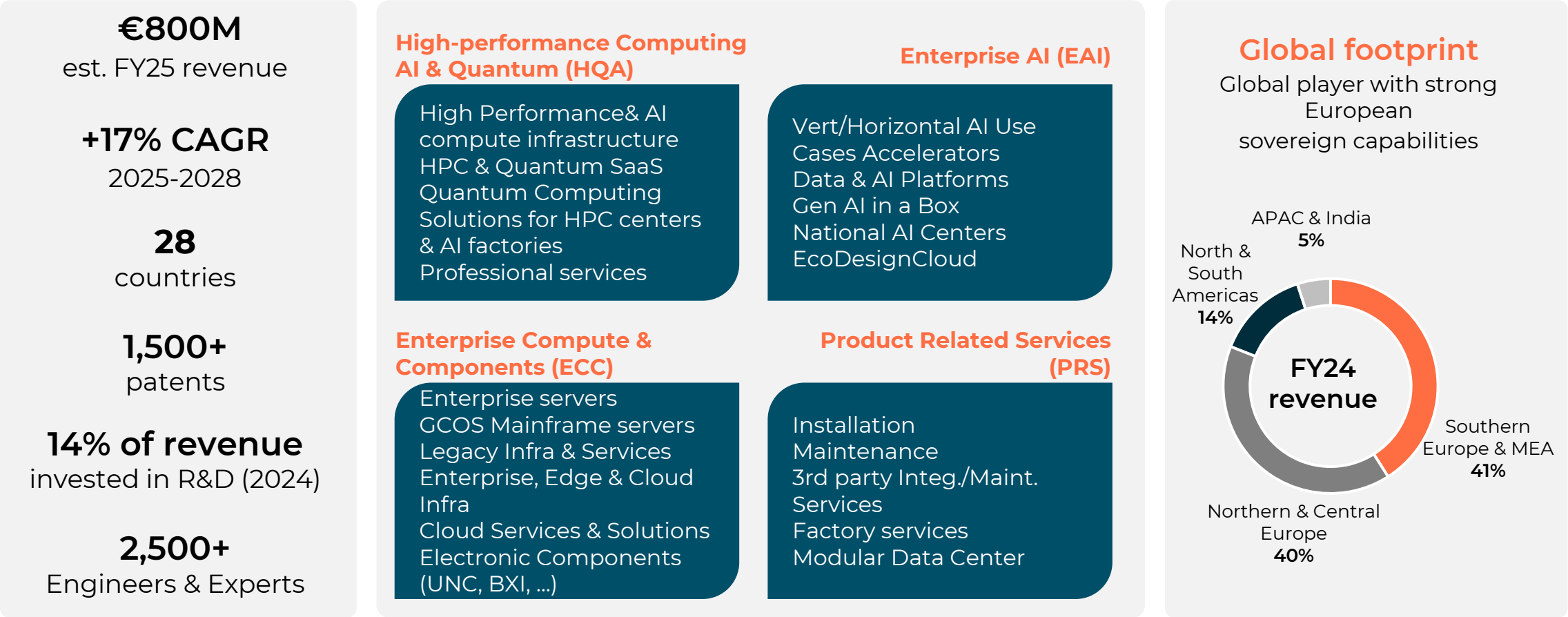
New Architectures for AI-Driven Performance Computing

Jean-Pierre Panziera
jean-pierre.panziera@eviden.com

EPI Forum
Paris, 6–7 October, 2025

PR: On July 31, 2025, Atos Group signed a share purchase agreement with the French State for the sale of its Advanced Computing business... to be finalized H1-2026

Advanced computing & AI: Powering Europe’s AI future



A global supply chain with a factory unique in Europe using advanced technologies to deliver 4 Exascale machines annually and up to 5 with its capabilities mirrored in India and Brazil

JUPITER – the Fastest Supercomputer in Europe

Solution and benefits

- The **fastest supercomputer in Europe**, #X4 TOP500 (June 2025)
- BullSequana XH3000 integrates **23,536** NVIDIA GH200 and **1.7M CPU cores**
- A total of 125 racks are housed in a Modular Data Center, **designed and delivered by Eviden in only 9 months**
- Modular Design: All 50 pre-built, interchangeable containers are **pre-integrated at our Angers facility** with cooling, power, networking, and cabling to streamline on-site deployment
- The **most energy-efficient HPC among the top 5 fastest supercomputers in the world**, empowered by the Eviden DLC
- Its early development module (JEDI) has achieved the **#1 position on the GREEN500 ranking three times in a row**
- **The 1st system**, which will integrate **European sovereign CPU – SiPearl Rhea**



793 petaflops → 1 exaflop

#1 Europe Top500

#4 Worldwide Top500

#1 JEDI Green500



All in one compute and modular data center

Trees don't grow to the sky ... but AI might (?)

Nature, 09/10/2024: **Chemistry Nobel** goes to developers of **AlphaFold** AI that predicts protein structures

European Union, 30/06/2025: Overwhelming response as **76** respondents express interest in the **European AI Gigafactories** initiative

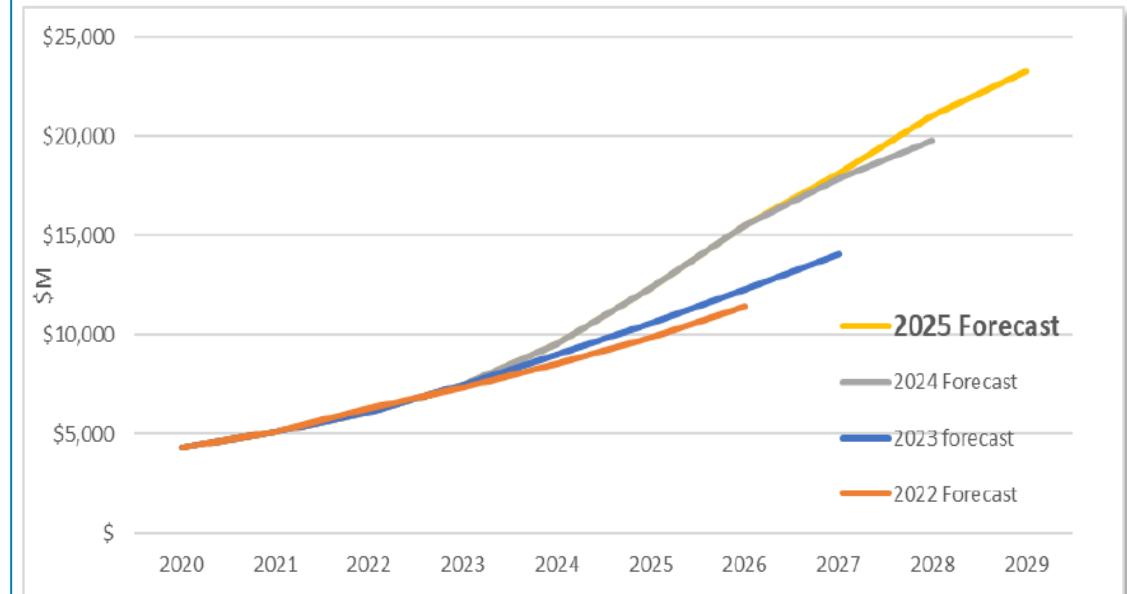
Generative AI:

ChatGPT, Claude, DeepSeek, Gemini,
Perplexity, Le Chat ...
TildeOpen LLM

AI infrastructure **>300B\$ 2025**.
1.5T\$ spending in AI in 2025, 2T\$ in 2026(Gartner)



HPC-AI Cloud Spending Forecast Comparison



Source: Hyperion Research, 2025

Trees don't grow to the sky

Forbes 30/09/2025: Why AI ROI Continues To Be Elusive Despite Broad Adoption

One of the most pressing challenges facing enterprises today is the **disconnect** between ubiquitous AI adoption at the individual level and the **absence of** transformational **business impact** at the organizational level

WSJ, 29/09/2025: Tech-Bubble Fears Rise in New Investor Poll

Concerns that U.S. technology stocks are in a **bubble crept higher** in Deutsche Bank's quarterly survey of institutional investors.

On a scale from 1 to 10, from "no bubble" to "extreme bubble", respondents gave tech stocks a **7.4** ...

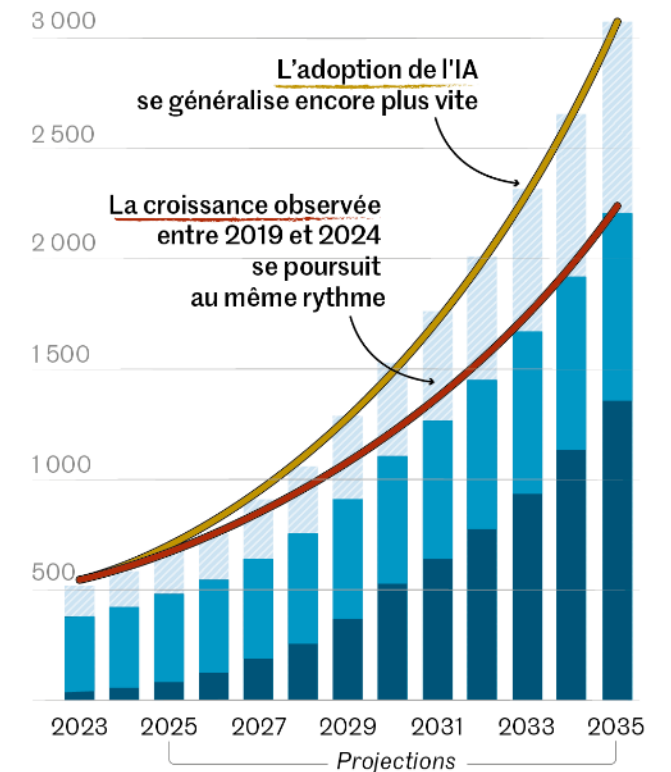
Le Monde, 02/10/2025: L'IA, dévoreuse d'énergie, pourrait entraîner une pénurie d'électricité aux Etats-Unis

la demande d'électricité aux Etats-Unis devrait augmenter de **25 % d'ici à 2030** et de **78 % d'ici à 2050**, par rapport à 2023, entraînant des hausses de prix de 15 % à 40 % d'ici à 2030

Une consommation électrique exponentielle

Consommation d'électricité des centres de données en TWh

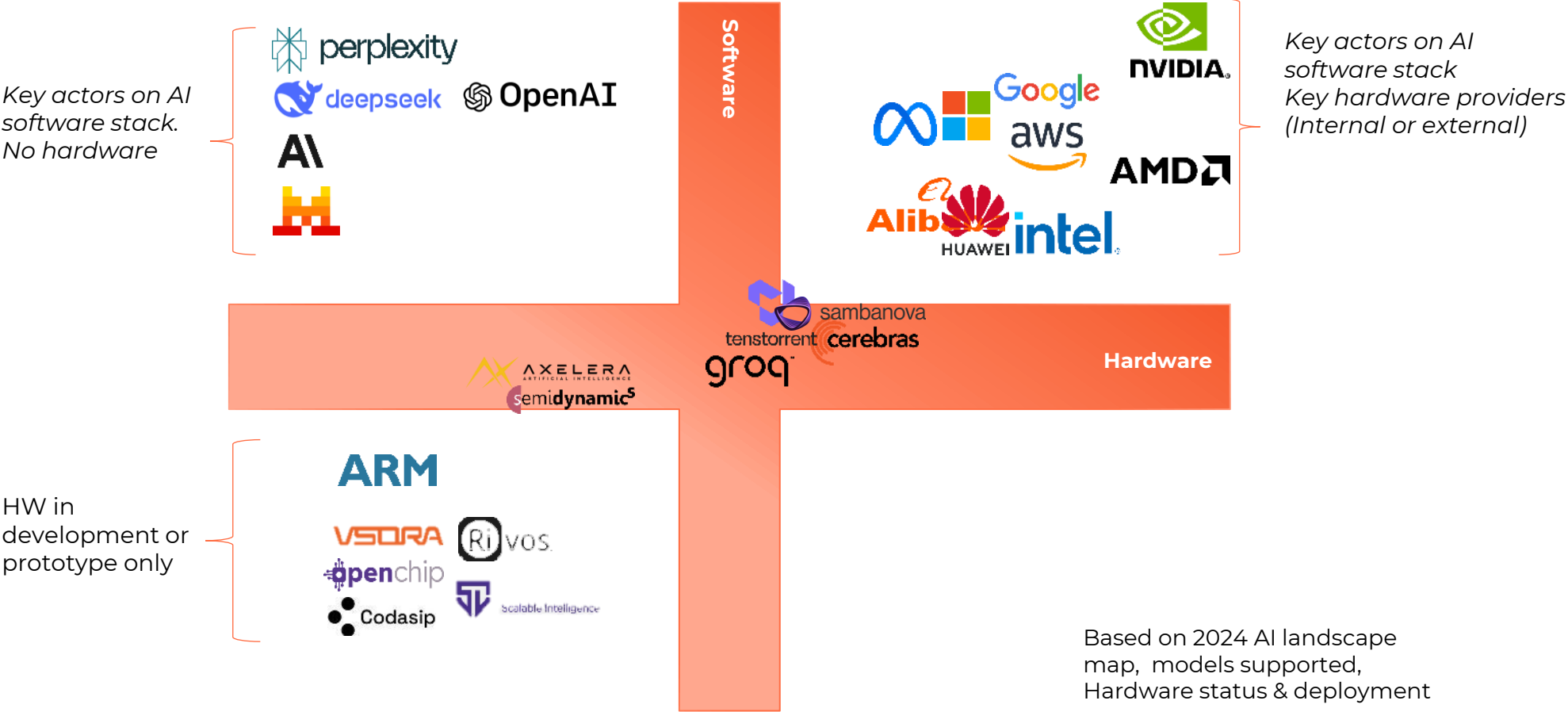
■ Intelligence artificielle (IA) générative ■ Consommation hors IA
■ Cryptomonnaies



Source : The Shift project • Infographie Le Monde

With AI new world, new players

A partial/subjective view



2030: We'll still need (high end) CPUs

For HPC, AI inference

Cores:

- 350 to 400 cores
- 1k small cores

Memory:

- 20+CH DDR6 (2.5-3TB/s per socket)
- SOCAMM2 LPDDR6

Performance:

- Enhance AI features for Inference
- FP64 TFlops: 25-40

IO:

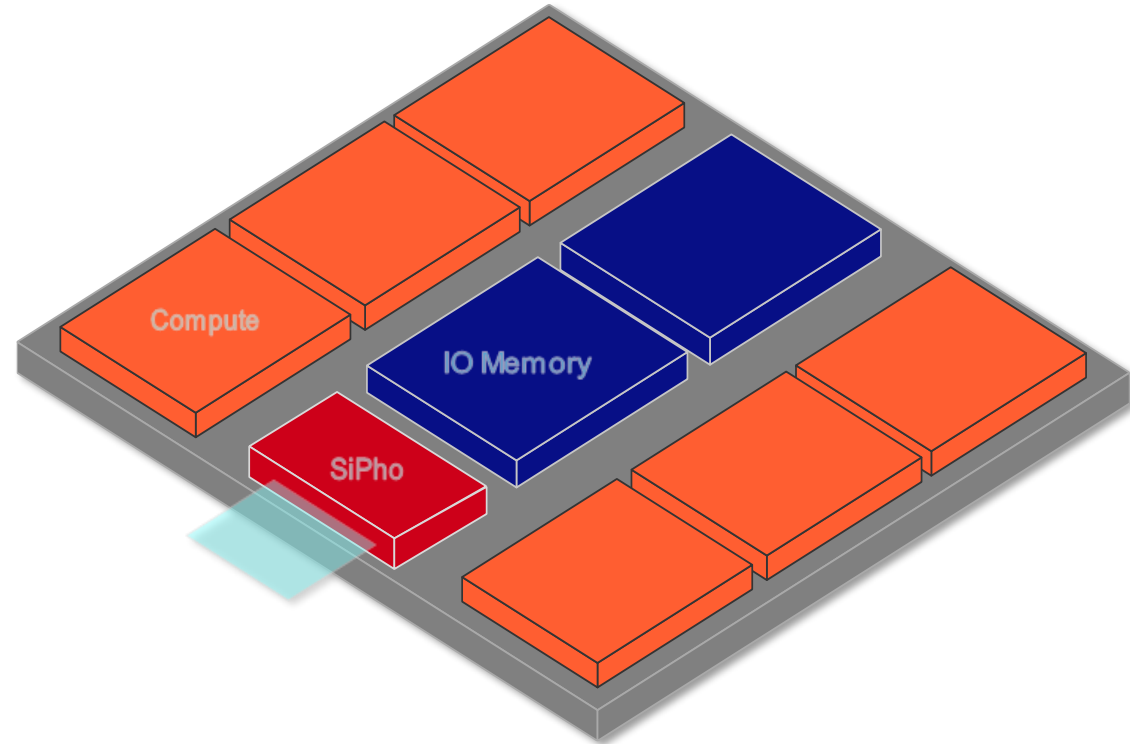
- 160-192L PCIe gen 7/8 (256-512GB/s)
- **CoPackage Silicon Photonics**

Power:

- 800W to **1200W** DLC
- <400W AC

Mechanical:

- (80-120)mm*(85-90)mm - 7000-10000mm²
- Manuf. Proc. A14-A10 – 14A, glass substrate



2030: will heavily use GPUs/Accelerators

HPC/inference

Key Features:

- FP64-FP32 Vector & Matrix Flops
- Large memory bandwidth

Memory:

- 12+ HBM5/6 – 1.5-2TB -> 30-80TB/s
- GDDR – HBM reserved for AI ???

Performance:

- FP64 TFlops: 250-500 
- AI features for Inference

IO:

- PCIe gen 7/8 (256-512GB/s)
- Direct link to the IO fabric
- Proprietary CPU to GPU link

Power:

- 4000-**8000W** DLC
- ~1000W AC

Mechanical :

- 100.000-150.000mm² module



trend towards low precision.

AI training

Key Features:

- FP32/16/8 matrix Flops
- Large memory bandwidth and capacity
- Scale-Up network

Memory:

- 12+ HBM5/6 – 1.5-2TB – 20-80TB/s

Performance:

- FP32/16/8 TFlops: 200-300/600-1200/1000-4500

IO:

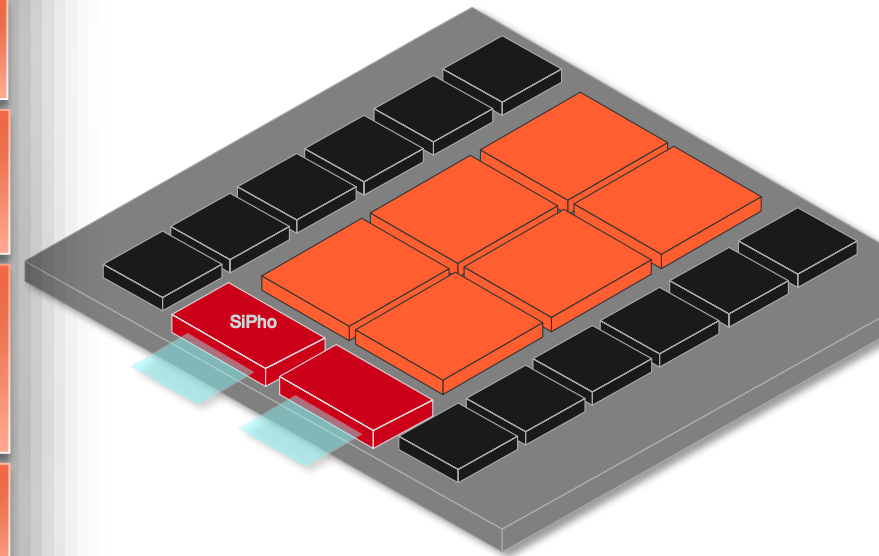
- Scaleup : Silicon Photonics integration
- Direct link to the IO fabric
- Proprietary CPU to GPU link

Power:

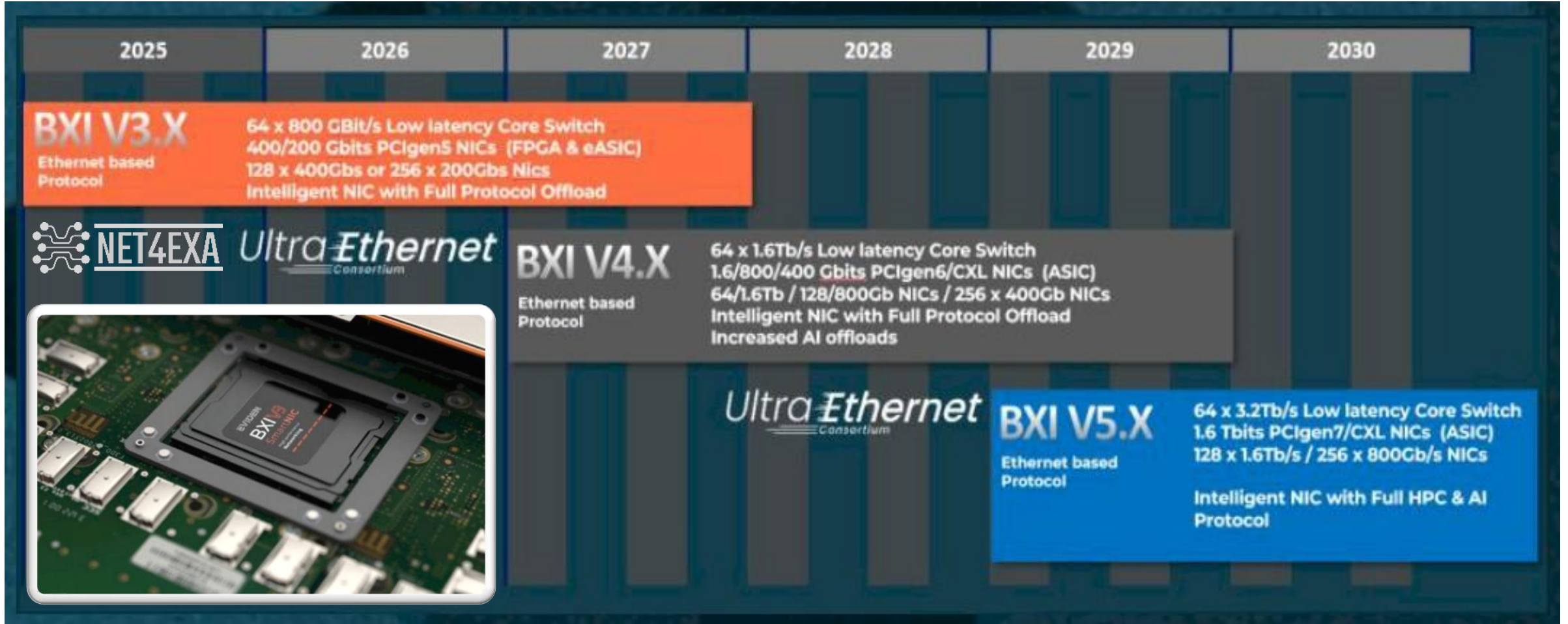
- 4000-**8000W** DLC

Mechanical:

- 100.000-150.000mm² module



BXI v3,4,5



Power them up, Cool them down



BullSequana X1000
(2016)

BullSequana XH2000
(2019)

BullSequana XH3000
(2023)

BullSequana XH3500
(2026)

2nd generation DLC

- 40°C rack inlet water temperature support
- **Increased density**
- Designed to better scale to large PetaScale systems
- Support of **InfiniBand EDR**

3rd generation DLC

- including **PSUs DLC**
- **“All-In-One” Rack** from single rack Exascale systems
- **87kW DC** (N+1)
- **Increased flexibility** of compute and interconnect supported
- Support of **InfiniBand HDR, BXI V2**

4th generation DLC

- **+70% of power and cooling**
- **147kW DC** (N+1)
- **Increased flexibility** of compute and interconnect supported
- Support of **InfiniBand NDR, BXI V2**

5th generation DLC

- **++% of power and cooling**
- **Increased flexibility** of compute and interconnect supported
- Support of **InfiniBand XDR, BXI V3**

Advanced HPC/AI computing in 2030

- **AI dominated**
 - **AI learning**
 - **AI inference**
 - **AI boosted HPC**
- **GPU/Accelerator driven**
... while still needing CPUs
- **Dense**
- **Powerful / Power hungry**
- **Baking **HOT**, Liquid Cool**



Thank you