

AI Infra at Meta

Wes Bland, Research Scientist
Meta AI

Meta's GPU Computing Infrastructure

Massive Infrastructure spanning dozens of data centers around the world

Several multi-GW data centers including the recently announced Prometheus in Ohio

AI is a significant part of our compute infrastructure, powered by multiple types of GPUs as well as our internal silicon

Just our NVIDIA GPU portfolio is over 600K H100 equivalents overall

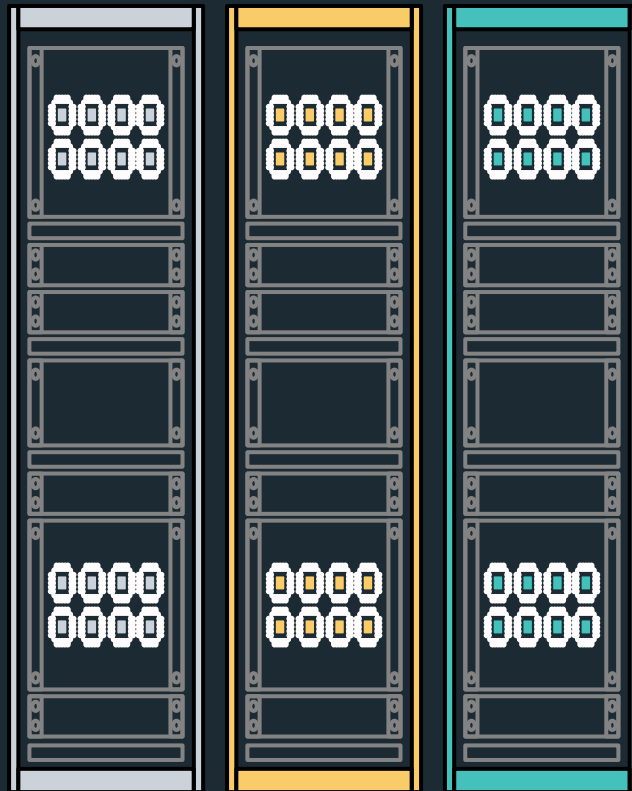
<https://engineering.fb.com/2024/03/12/data-center-engineering/building-metas-genai-infrastructure/>



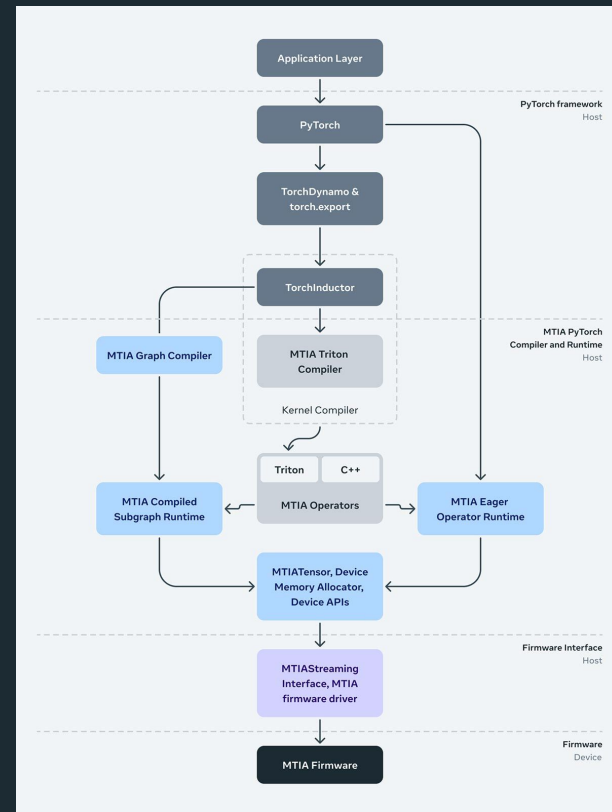
Illustration: Nvidia HGX board design
(<https://nvidianews.nvidia.com/news/nvidia-hpc-platform-hopper-quantum-2-worldwide-adoption>)

Hardware Heterogeneity

- Supply and cost constraints can push training towards heterogeneous systems
- Ability to train models on a mixed variety of hardware types, either from the same vendor or different vendors is critical
- Numerics, performance, and reliability in such mixed environments are important to consider



Meta's MTIA in-house silicon for performance and cost



Data and Storage

High quality and high volume data are the backbone of deep learning

Llama 4 was trained on 30T tokens of data - 2x larger than Llama 3 - using a mix of publicly available online data

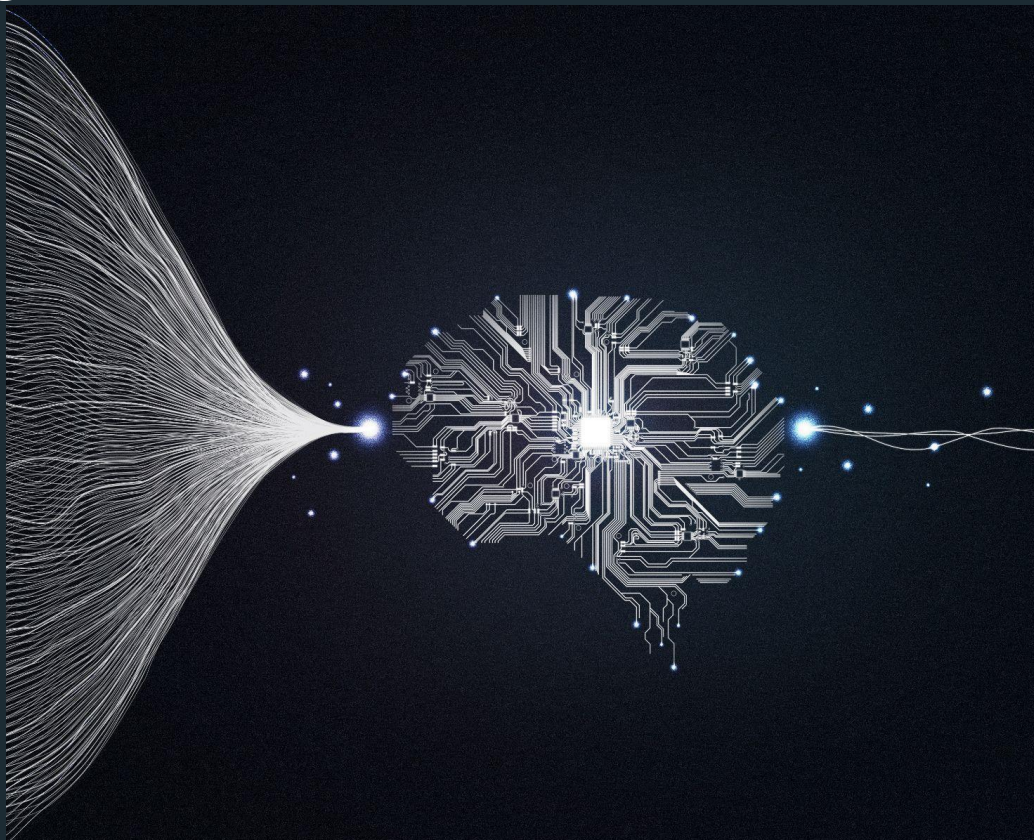
Storage infrastructure based on flash media + Meta's 'Tectonic' distributed storage solution

Also co-developed NFS solution with Hammerspace to allow for interactive debugging

<https://engineering.fb.com/2024/03/12/data-center-engineering/building-metas-genai-infrastructure/>

<https://ai.meta.com/blog/llama-4-multimodal-intelligence/>

<https://www.usenix.org/conference/fast21/presentation/pan>



Network Traffic for AI Training

- **Challenges**

- Elephant flows: few extremely large continuous flows
- Low entropy: less variation, more likely to cause hash collisions
- Oscillatory behavior during congestion

- **Solution: DSF**

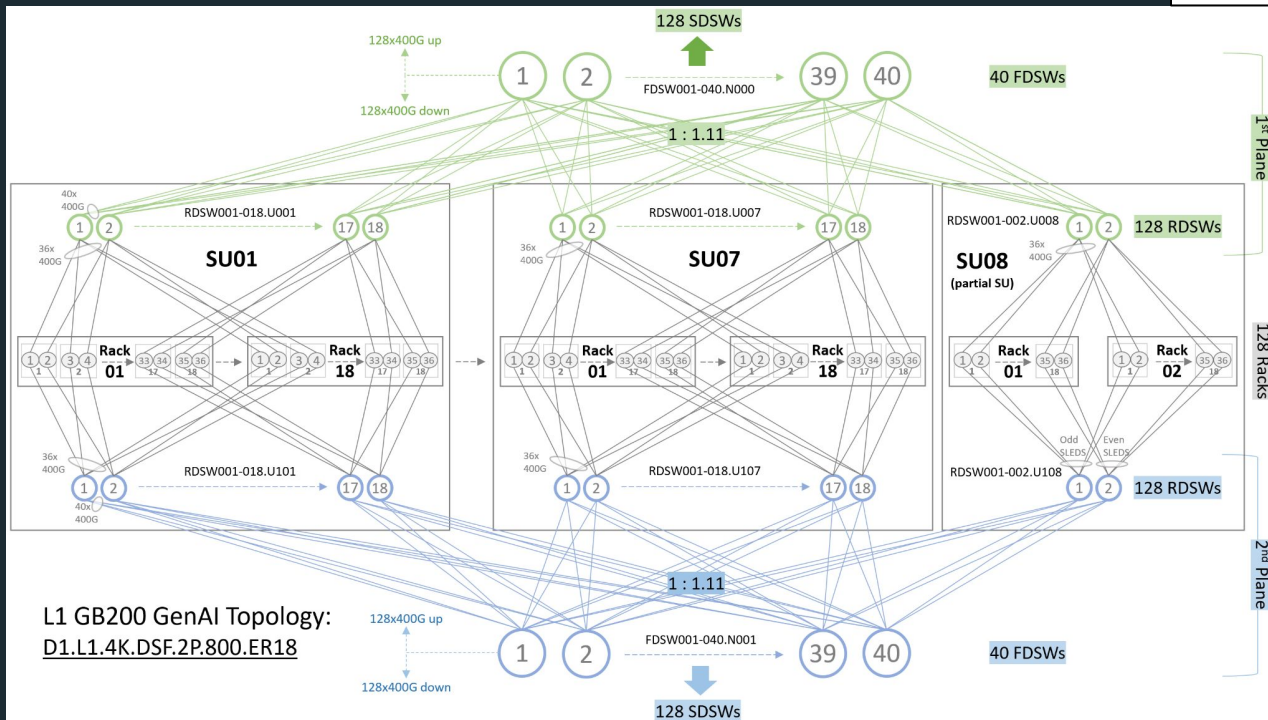
- Disaggregated **S**cheduled **F**abric
- Near-optimal load balancing
- Smoother bandwidth delivery: credit allocation
- Flexibility/optionality for endpoints: fabric performs spray/reassembly

DSF: Disaggregated Scheduled Fabric

RDSW = Disaggregated Line Card

FDSW = Disaggregated Fabric Card

SDSW = 2nd stage Fabric Card



L1 GB200 GenAI Topology:
D1.L1.4K.DSF.2P.800.ER18

Dual-Stage Topology announced at OCP APAC Summit, August 2025

Switch Hops, Physical Distances, Latency



Meta's Mesa Data Center in Arizona

Power constraints often force us to spread out GPUs over large physical distances

Makes adaptive routing challenging:

- Large window of out-of-order packets (high network bandwidth-delay product)

- Need custom transports with more flexible software than IB verbs

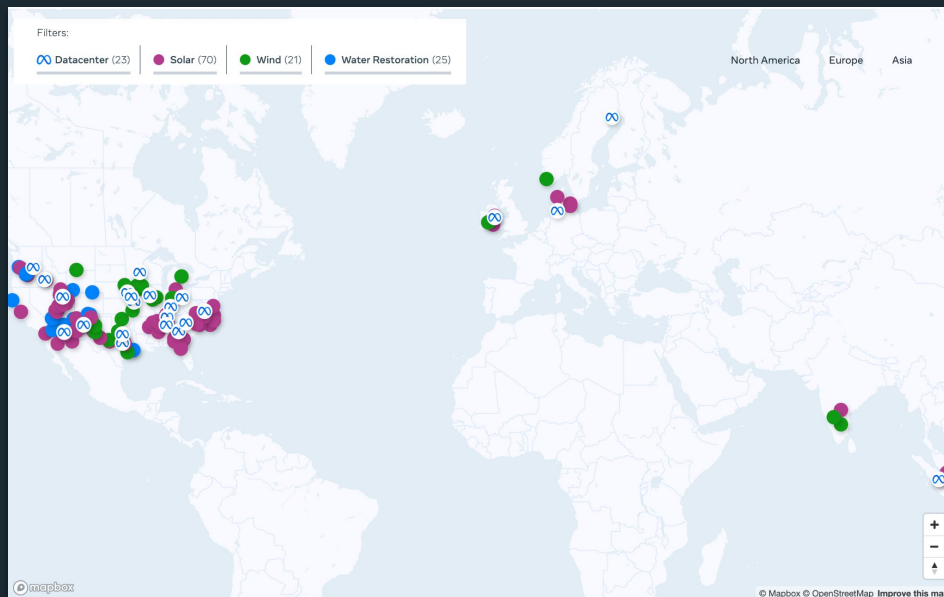
Power is a Key Basic Resource

Because Meta is committed to sustainable data centers (100% renewable energy), infrastructure has to be very power efficient

Hardware/software systems need to be heavily optimized for perf/watt, rather than raw perf

- Optics vs Copper
- More GPUs, each at lower power
- Meta internal silicon for aggressive perf/watt optimization

<https://sustainability.fb.com/data-centers/>



Parting Thoughts

Scale changes everything

- Raw performance is almost never the core problem
- Performance is often constrained by power, physical layout, system reliability, and other physical constraints

As we continue to scale, a new generation of problems is upon us

- Compute cost, reliability, data, power
- Need the community's help in new innovative solutions for these problems

Opportunities to Work Together

Full-time research scientists and software engineers in Computer Science, specifically with experience in architecture, networking, communications, HPC, and AI/ML systems

Internship and sabbatical opportunities for 2026

Research collaboration opportunities

Get Involved!

Email: wbland@meta.com

Open Positions:

https://fb.me/hot_interconnects_2025

