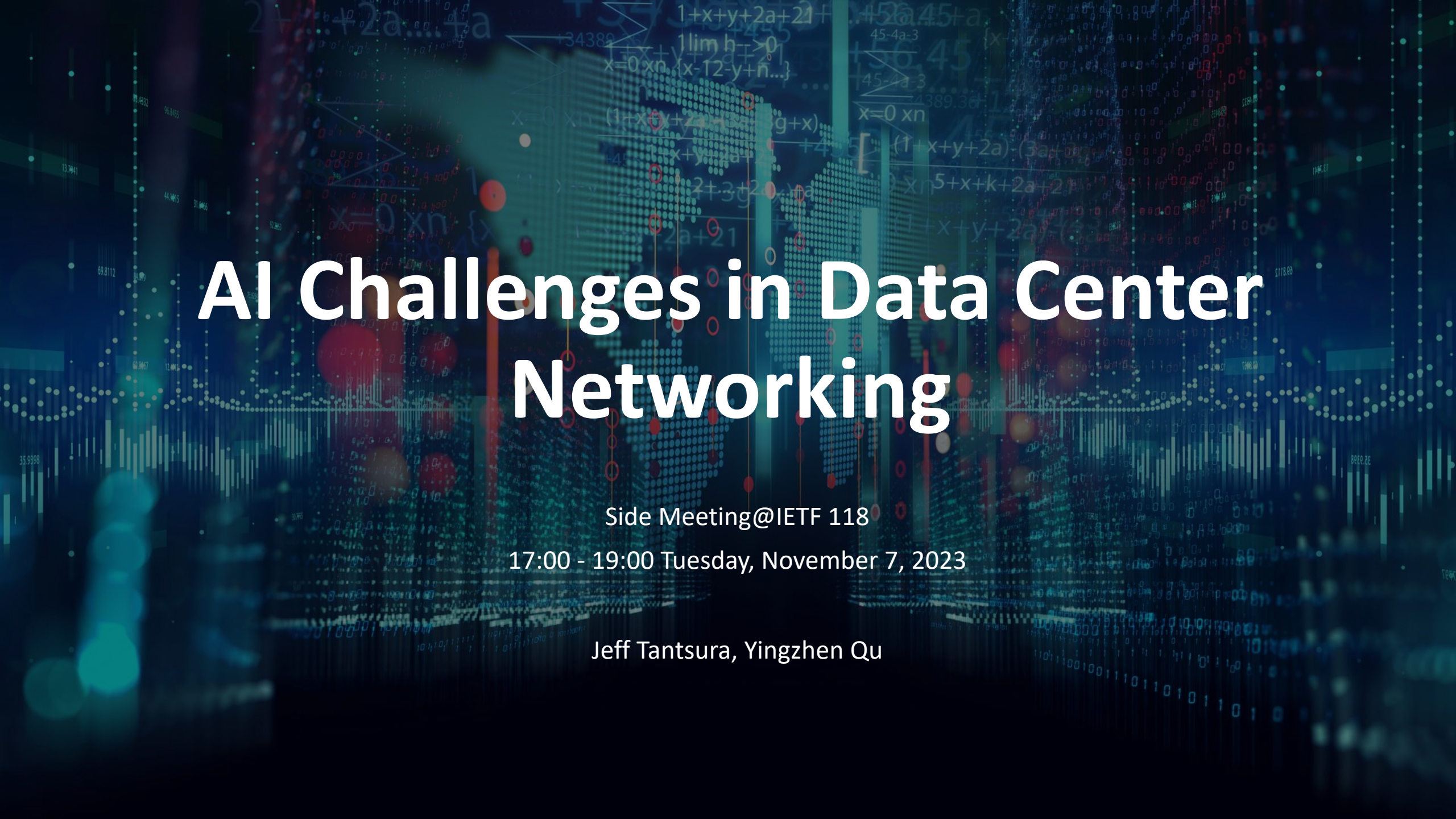


AI Challenges in Data Center Networking

Side Meeting@IETF 118

17:00 - 19:00 Tuesday, November 7, 2023

Jeff Tantsura, Yingzhen Qu

Note Well

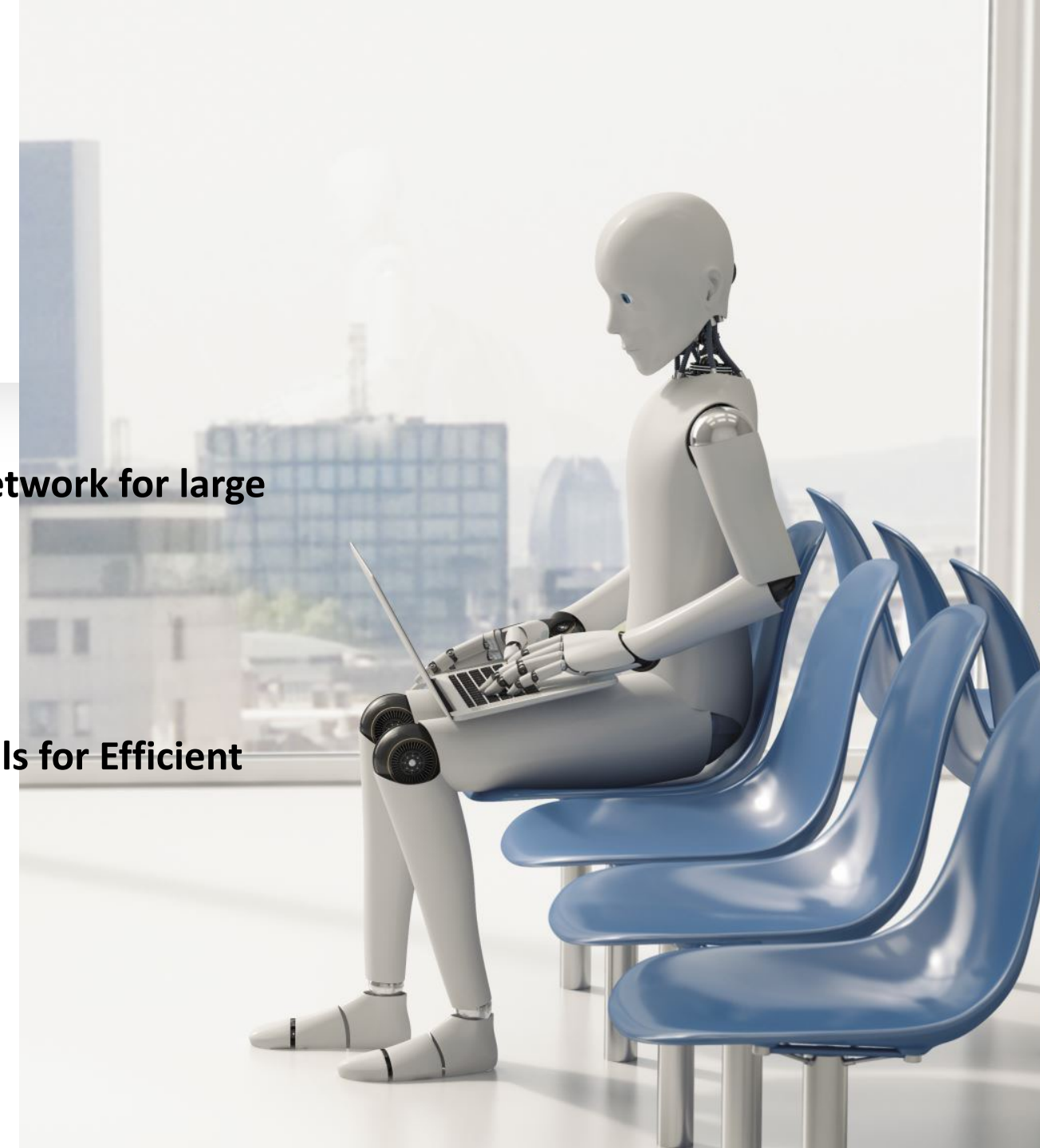
- This is a reminder of IETF policies in effect on various topics such as patents or code of conduct. It is only meant to point you in the right direction. Exceptions may apply. The IETF's patent policy and the definition of an IETF "contribution" and "participation" are set forth in BCP 79; please read it carefully.
- As a reminder:
  - By participating in the IETF, you agree to follow IETF processes and policies.
  - If you are aware that any IETF contribution is covered by patents or patent applications that are owned or controlled by you or your sponsor, you must disclose that fact, or not participate in the discussion.
  - As a participant in or attendee to any IETF activity you acknowledge that written, audio, video, and photographic records of meetings may be made public.
  - Personal information that you provide to IETF will be handled in accordance with the IETF Privacy Statement.
  - As a participant or attendee, you agree to work respectfully with other participants; please contact the ombudsteam (https://www.ietf.org/contact/ombudsteam/) if you have questions or concerns about this.
-

Definitive information is in the documents listed below and other IETF BCPs. For advice, please talk to WG chairs or ADs:

- BCP 9 (Internet Standards Process)
- BCP 25 (Working Group processes)
- BCP 25 (Anti-Harassment Procedures)
- BCP 54 (Code of Conduct)
- BCP 78 (Copyright)
- BCP 79 (Patents, Participation)
- https://www.ietf.org/privacy-policy/(Privacy Policy)

Agenda

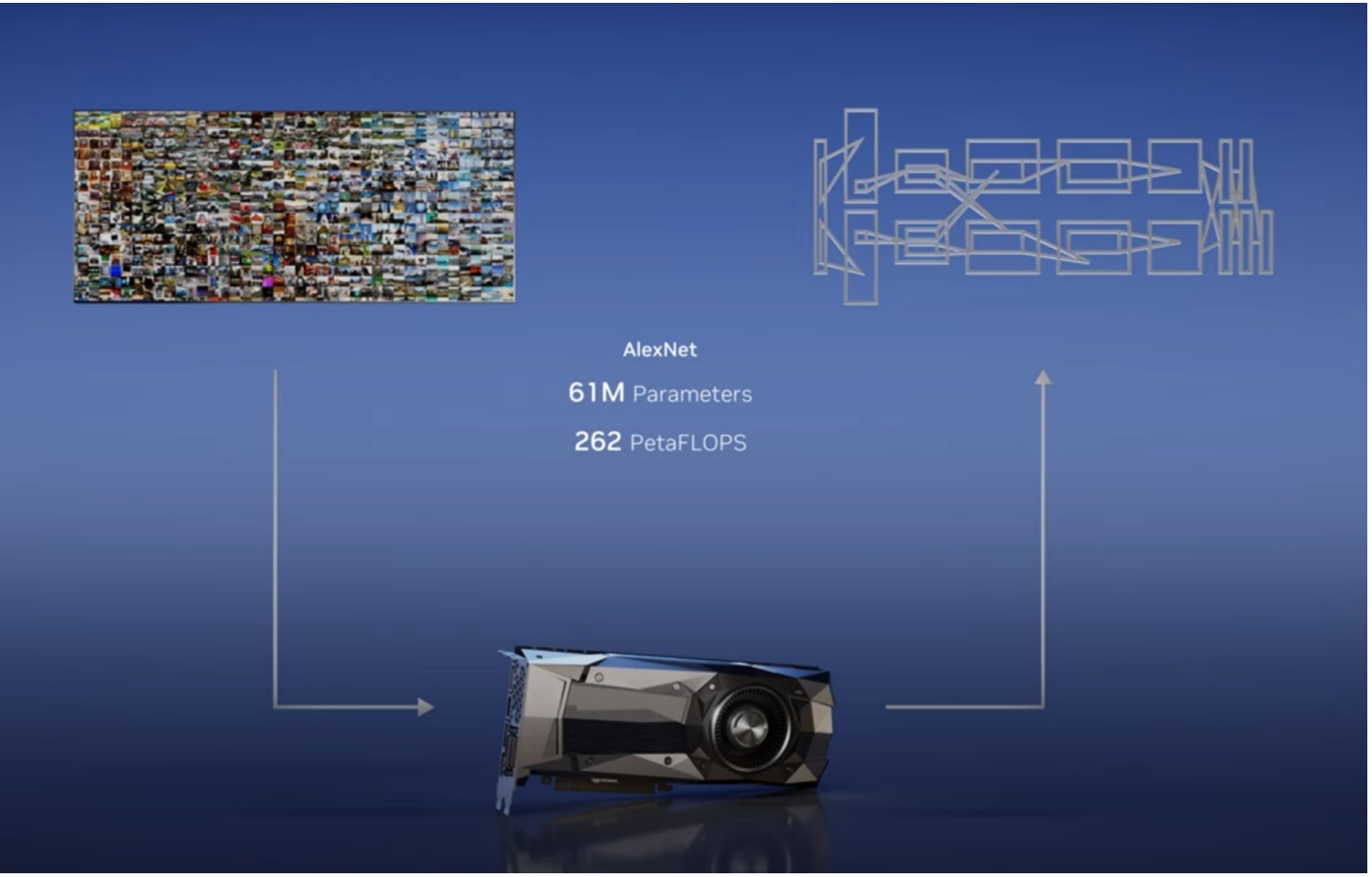
- **Chairs**
- **Networking in AI Clusters**
  Omer Shabtai (Nvidia)
- **Astral-Network: efficient large-scale datacenter network for large language model training**
  Baojia Li (Tencent)
- **Self-Adjusting Networks**
  Stefan Schmid (TU Berlin)
- **CSIG - Simple and Effective In-band Network Signals for Efficient Traffic Management in Datacenter Networks**
  Abhiram Ravi (Google)
- Open Discussions



AI networking – what is different?

- Performance optimized vs cost optimized (RDMA semantics)
  - cost of bit lost
  - cost of bit delayed
- JCT drives the networking
  - throughput != goodput != JCT
- Under-subscription is not an option – 90%+ utilization is mandatory for network bound jobs
  - every kW spent on networking is not spent on GPUs
- Many different networks
- Platform play -> communication libraries + smartNICs + switches

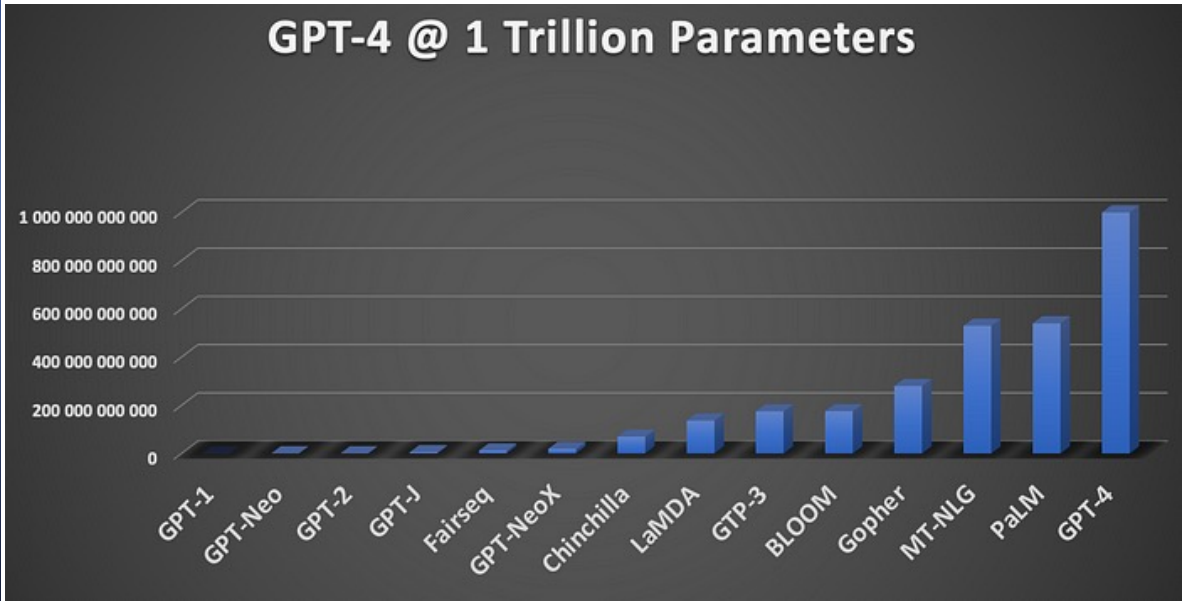
2012 – AlexNet – 2 GPUs/61M parameters



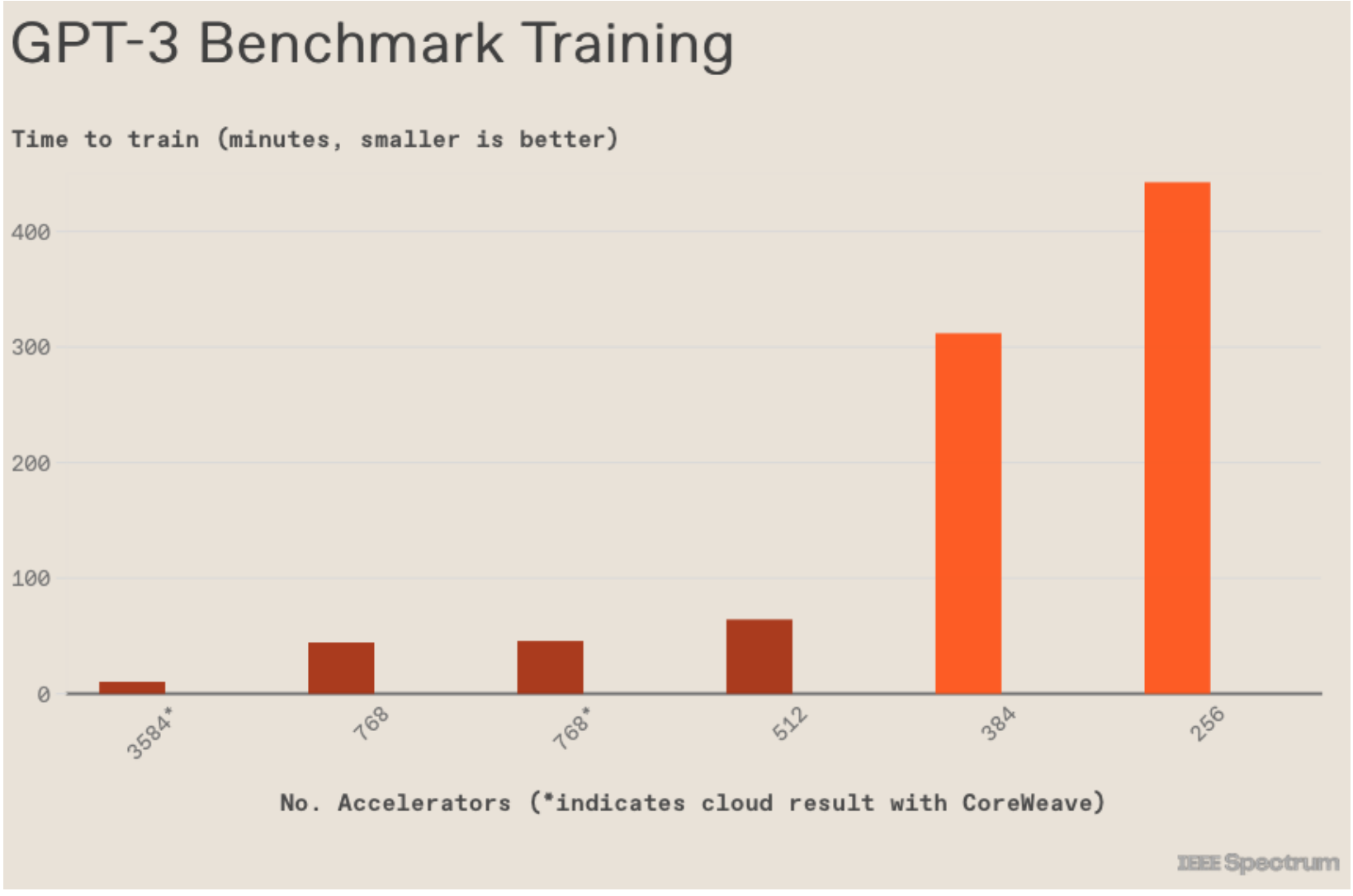
AlexNet was trained on a GTX 580 GPU with only 3 GB of memory which couldn't fit the entire network.

The network was split across 2 GPUs, with half of the neurons(feature maps) on each GPU

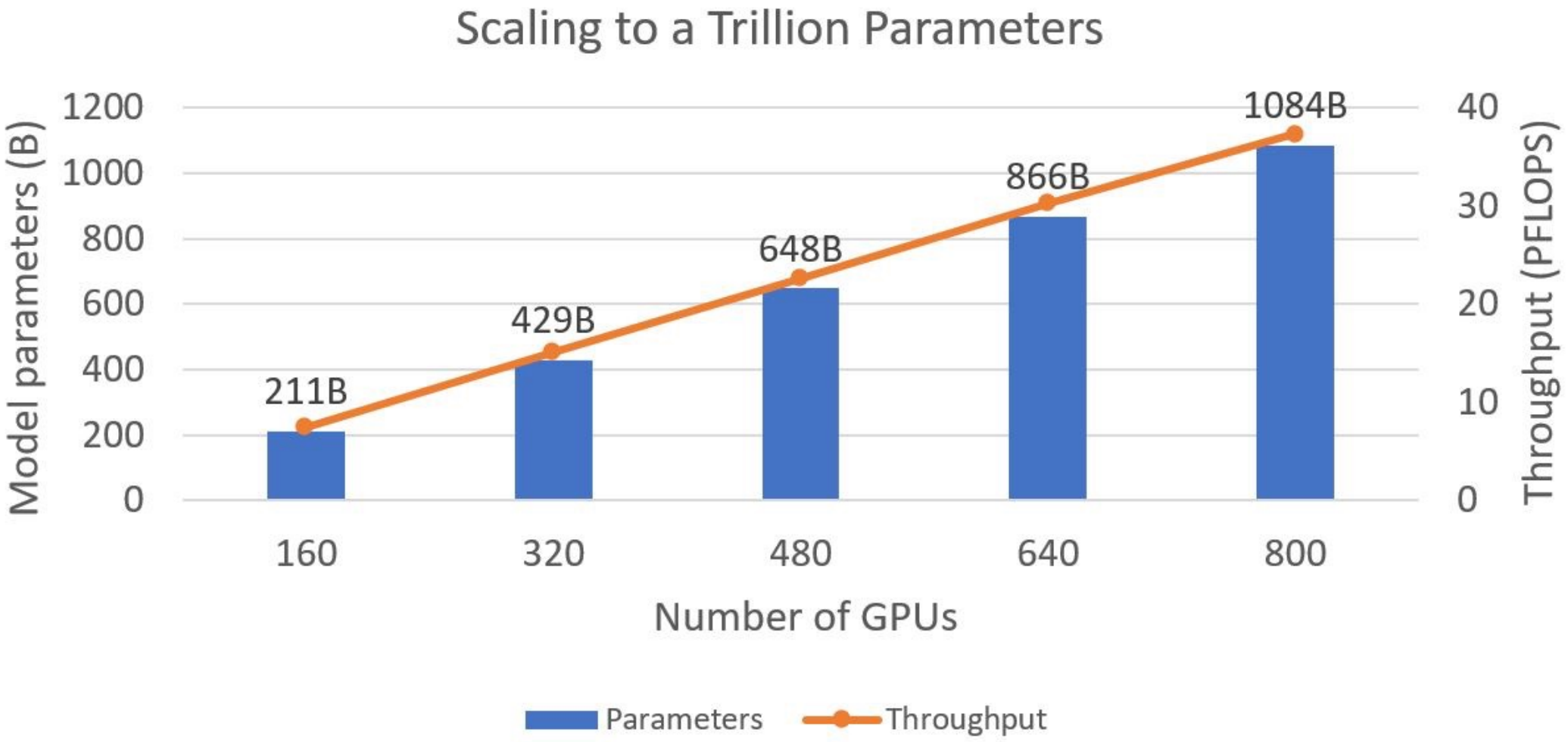
2022 – 512GPUs/175B; 2023 – 100T ->6K+ GPUs



Size does matter!



What drives the network size? Model size

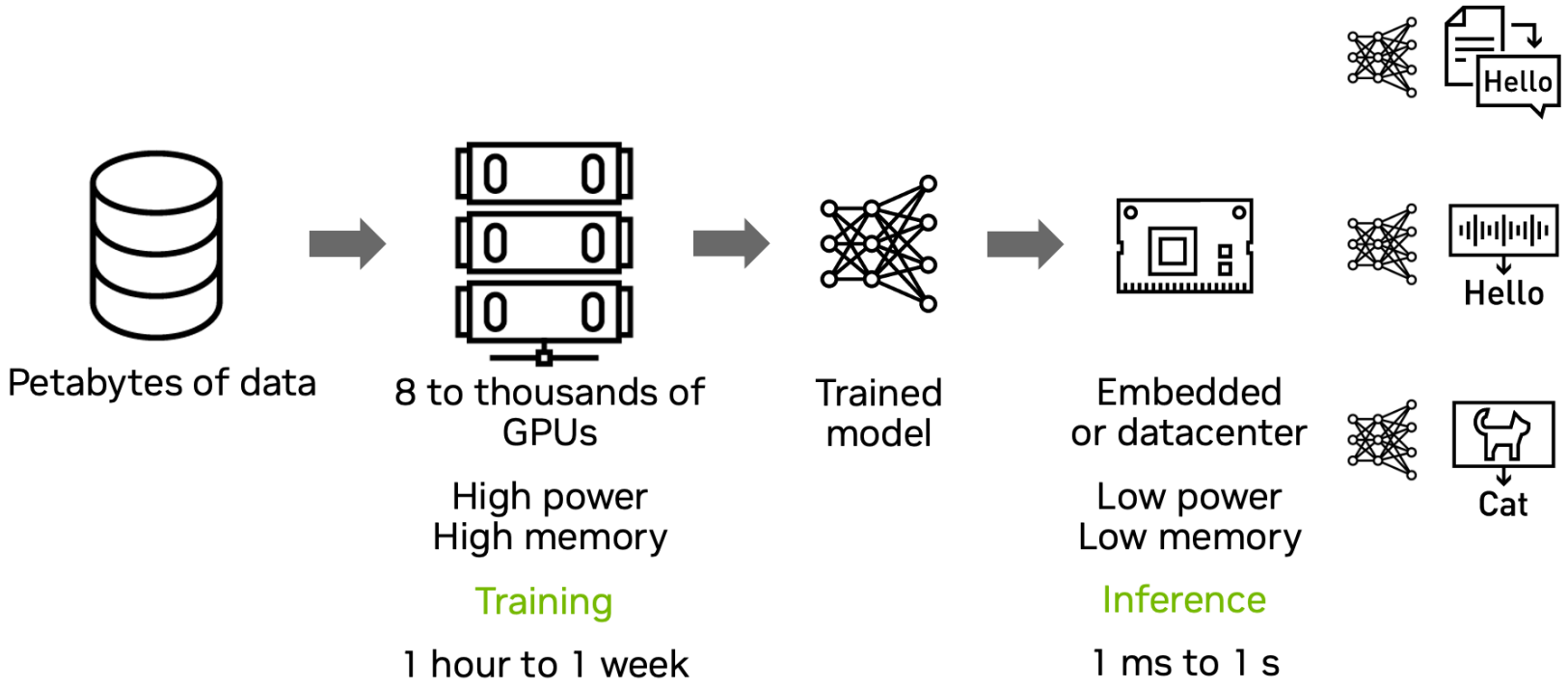


https://www.microsoft.com/en-us/research/blog/deepspeed-extreme-scale-model-training-for-everyone/

What drives the network size? Model size
META recent models(networking @scale 2023)

MODEL NAME	RELEASE DATE	MODEL SIZE	DATASET SIZE	TRAINING ZETA (1E21) FLOPS	TRAINING HW (COMPUTE)	TRAINING HW (NETWORK)	GPU HOURS (# GPUS X HOURS)
OPT	May 2022	175 B	300 B	430	1K A100	IB 200Gbps per GPU 25.6 TB/s bisection BW	800K
LLaMA	Feb 2023	65 B	1.4 T	600	2K A100	IB 200Gbps per GPU 51.2 TB/s bisection BW	1M
LLaMA2	July 2023	34 B	2 T	400	2K A100	RoCE 200Gbps per GPU 51.2 TB/s bisection BW	1M
LLaMA2	July 2023	70 B	2 T	800	2K A100	IB 200Gbps per GPU 51.2 TB/s bisection BW	1.7M

AI training/inference workflow



AI – how do we scale?

- To scale the model size -> parallelization is mandatory
  - Data
  - Model
  - Tensor

Meeting Materials

AIDC-IETF118: Meeting materials (github.com)