

Advancing Technologies for AI/ML Clusters and High- Performance Data Centers

Side Meeting@IETF 117

15:30 - 17:00 Monday, July 24, 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)

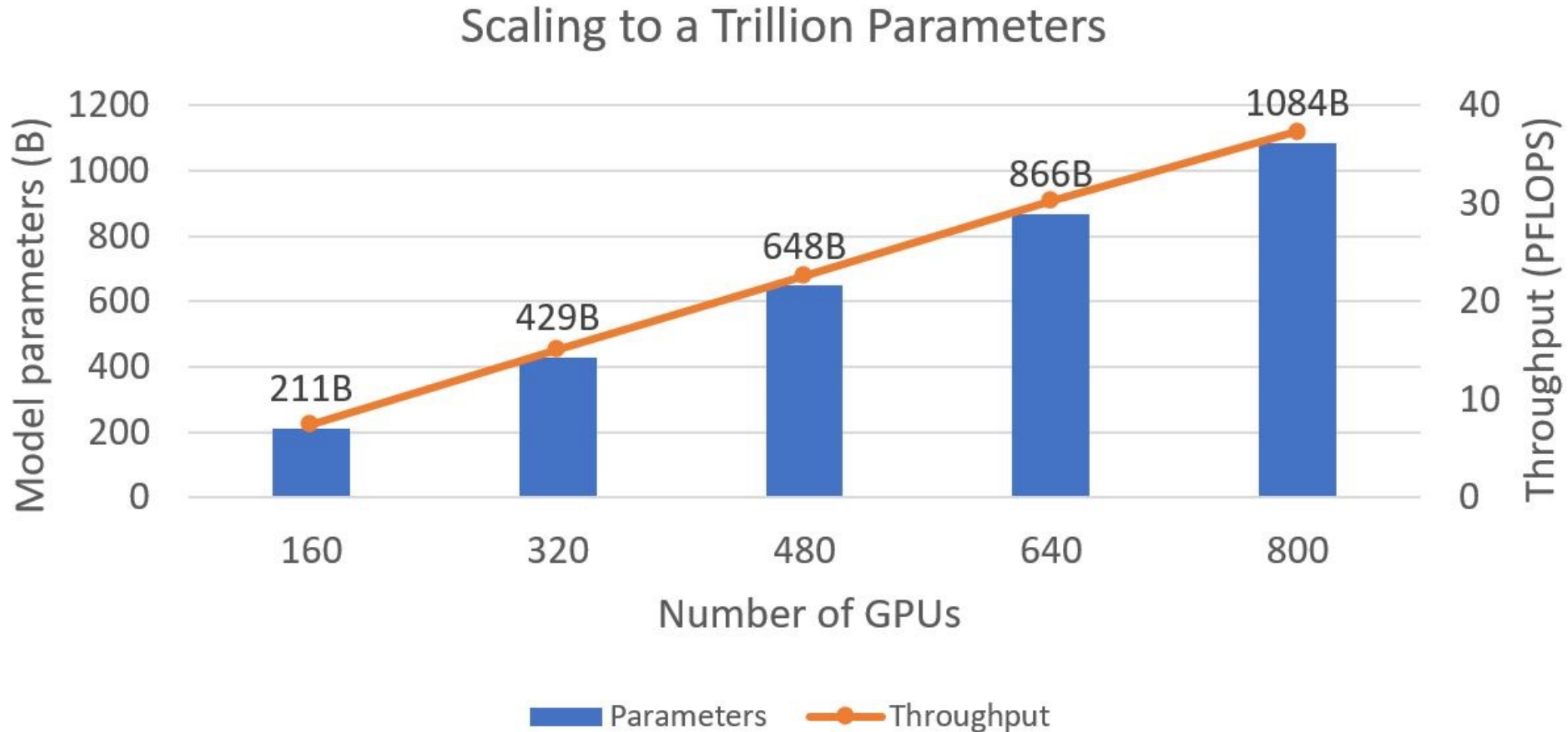
Agenda

- **Chairs**
- **Networking in AI/ML Cluster**
Barak Gafni (Nvidia) (20 mins)
- **Routing in Dragonfly Topology**
Dmitry Afanasiev (20 mins)
- **Aster!Network: efficient large-scale datacenter network for AI/ML**
Baojia Li (Tencent) (10 mins)
- **DC Routing**
Tony P (Juniper) (10 mins)
- **New Requirements and Thoughts for AI Data Center Networks - From a service provider's perspective**
Weiqiang Cheng (China Mobile) (10 mins)
- **Open Discussions (15 mins)**

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**

What drives the network size? Model size (params)



Meeting Materials

[GitHub Repository: AIDC-IETF117](#)