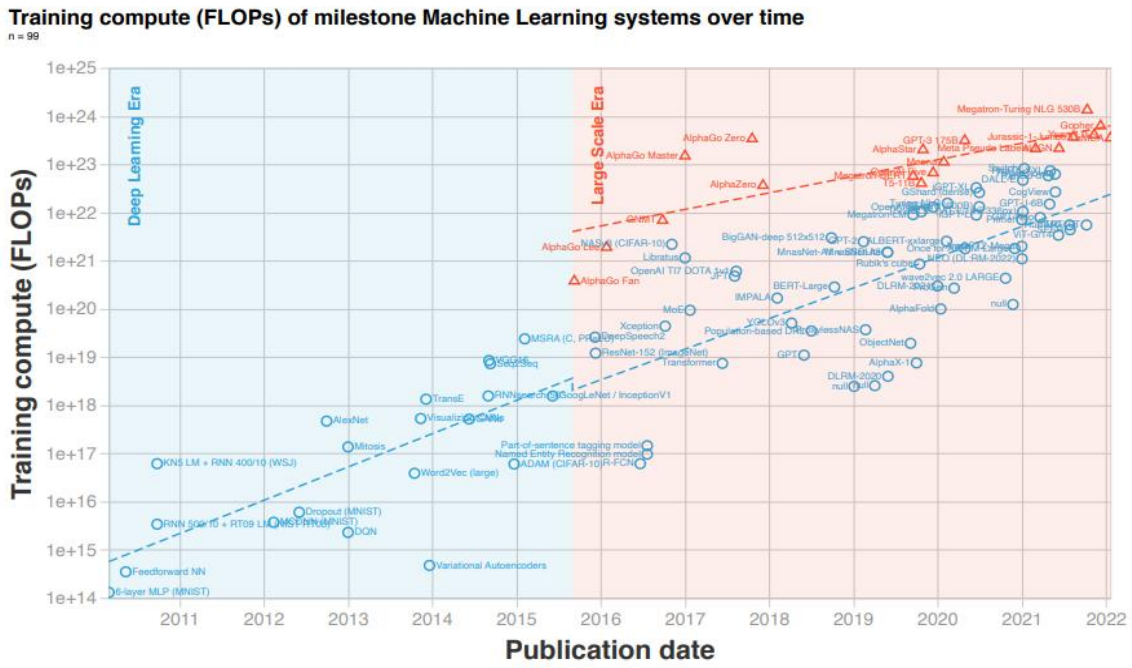
Requirements for AI Fabric

Weiqiang Cheng（China Mobile）

Ruixue Wang（China Mobile）

Rise of Large-Scale AI Models



Large-scale AI models represent a bifurcation of the Deep Learning Era (regular-scale models) trend.

- Large-scale AI models started with AlphaGo in late 2015
- The training compute is significantly higher than previous models.
- Doubling time of large-scale model is roughly every 10 months, much faster than Moore's law (roughly every 2 years)

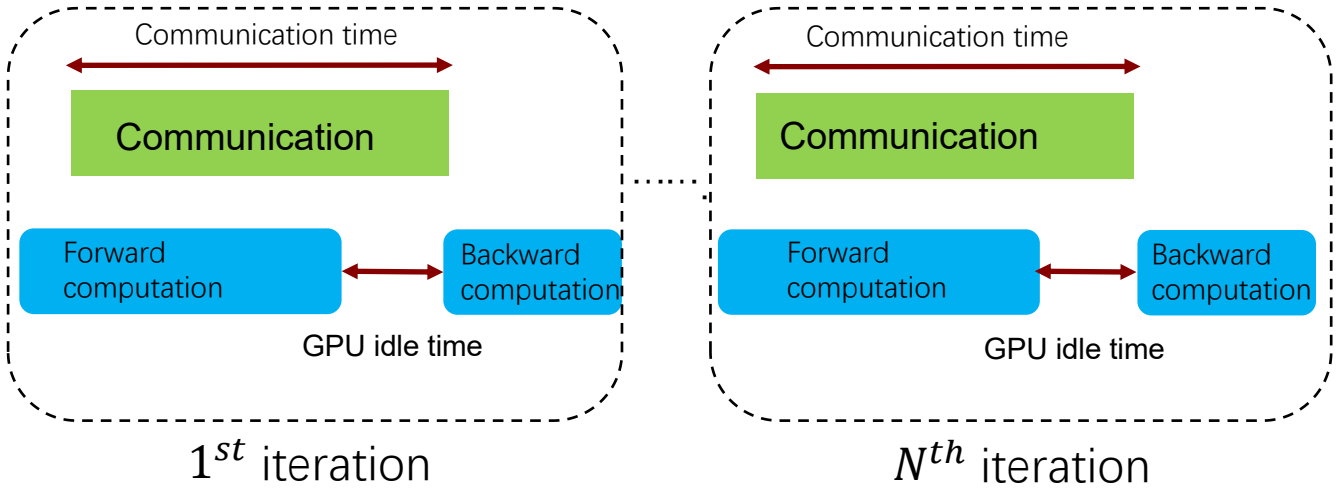
Source: 2022 International Joint Conference on Neural Networks (IJCNN), 2022, "Compute Trends Across Three Eras of Machine Learning"

ChatGPT draws strong attention to LLM

" We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. " -- OpenAI

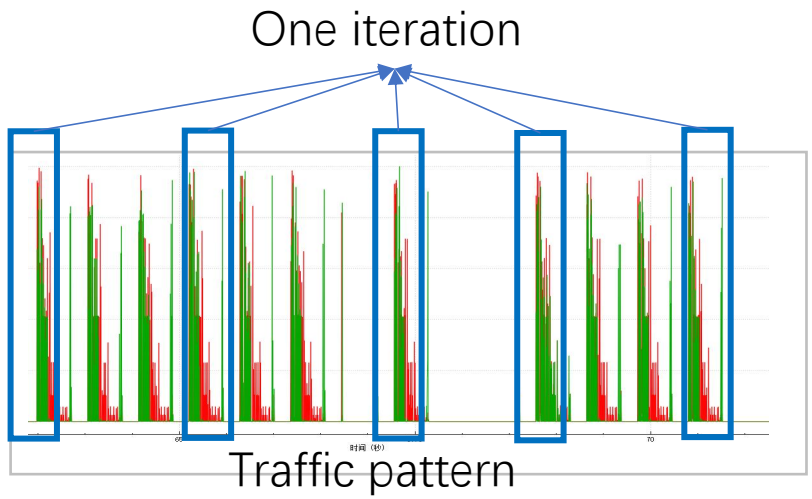
- ChatGPT is a Large Language Model (LLM) based on GPT-3.5 series.
- ChatGPT spreads to over one million users in 5 days from its release in 2022.

Communications in Large Models (1/2)



Periodical communication

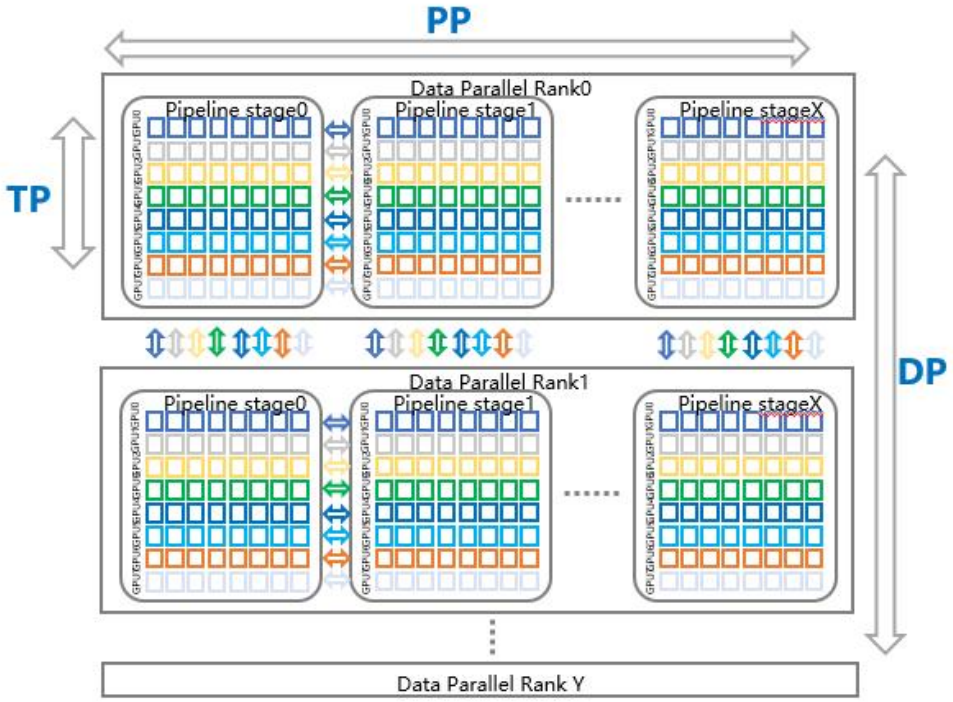
- AI training procedure consists of computing and communication. The computation tasks and communication are carried out periodically

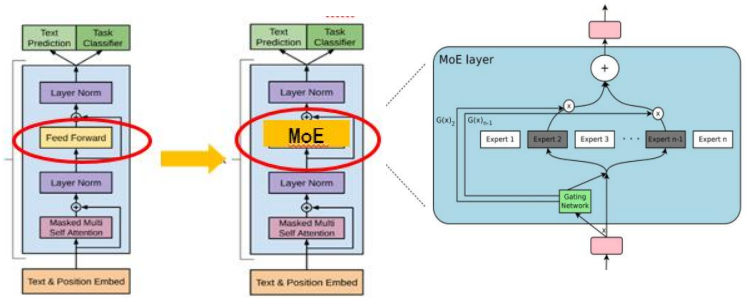


Burst traffic

- Communication remains consistent in each iteration, showing an ‘on-off’ type of burst traffic pattern

Communications in Large Models (2/2)





High bandwidth required with a small number of flows

Data parallelism: The training data is split into multiple mini-batches and trained in parallel on multiple AI chips

- Inter nodes communication (AllReduce)
- GB level traffic
- Bandwidth requirement: ★ ★ ★ ★

Tensor parallelism: Split the model into multiple sub-layers to run on multiple AI chips

- Intra node communication (AllReduce)
- 100s GB level traffic
- Bandwidth requirement: ★ ★ ★ ★ ★

Pipeline parallelism: Different layers of the model run on different AI chips

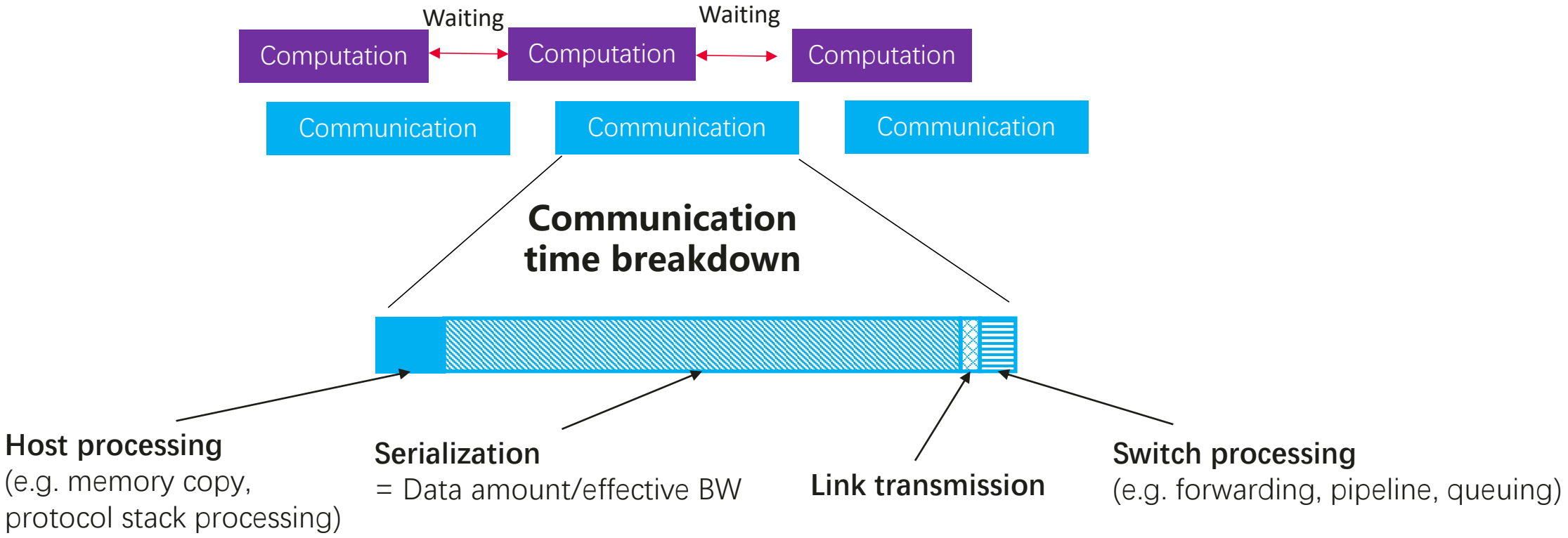
- Inter nodes communication (send/recv)
- 100s MB ~ GB level traffic
- Bandwidth requirement: ★ ★ ★

MOE: sparsely-gated Mixture-of Experts layer replaces FFN to efficiently expand large model parameters

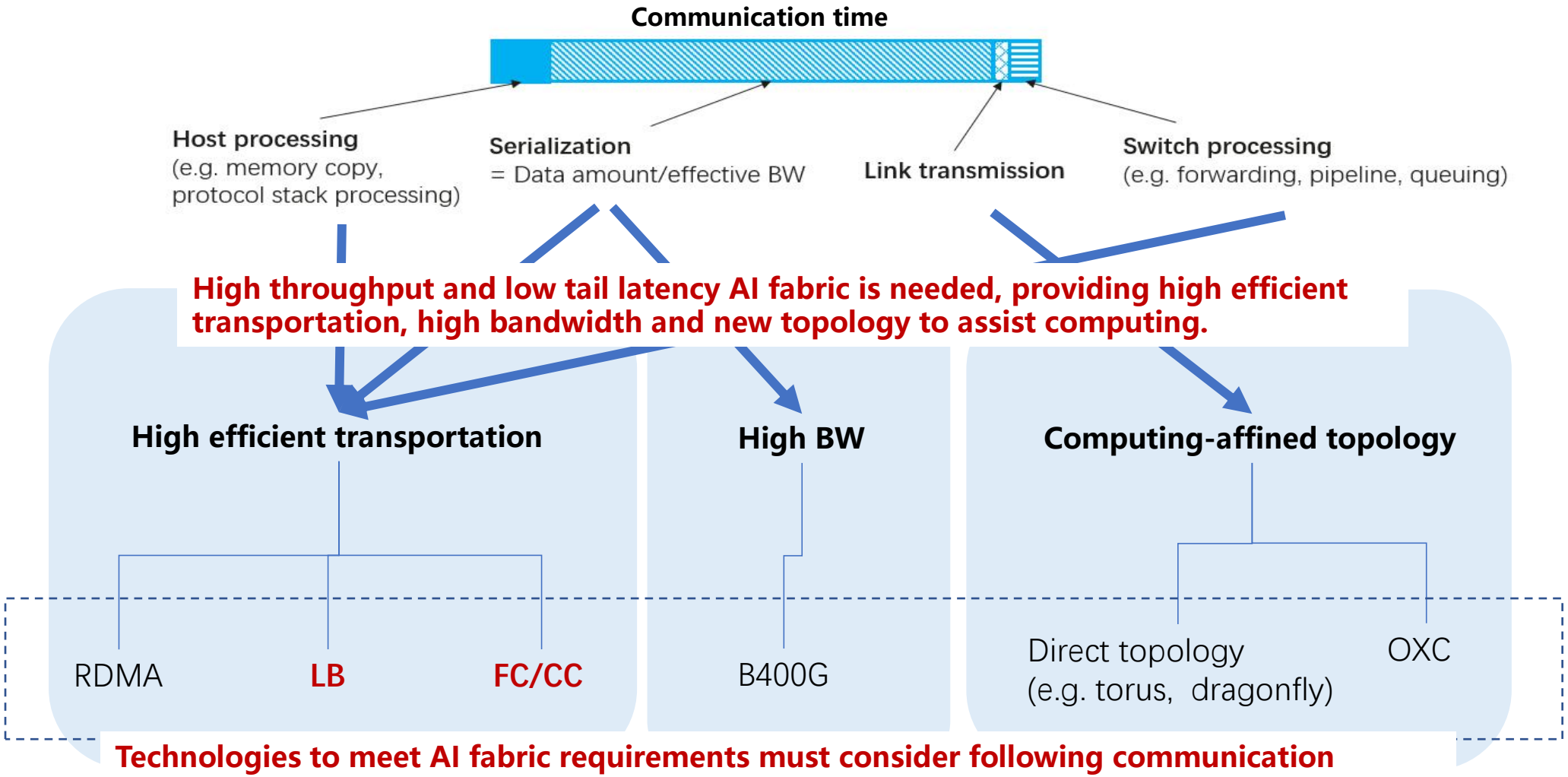
- Inter nodes communication (AlltoAll/AllReduce)
- GB level traffic
- Bandwidth requirement: ★ ★ ★ ★

Shortening Communication Time is Key to LM Training

- The completion of communication between all devices in each iteration pushes the computation tasks moving forward.
- Reducing communication time and overhead reduces GPU waiting time, thereby increasing computation efficiency.



Requirements of AI Fabric to Shorten Communication Time



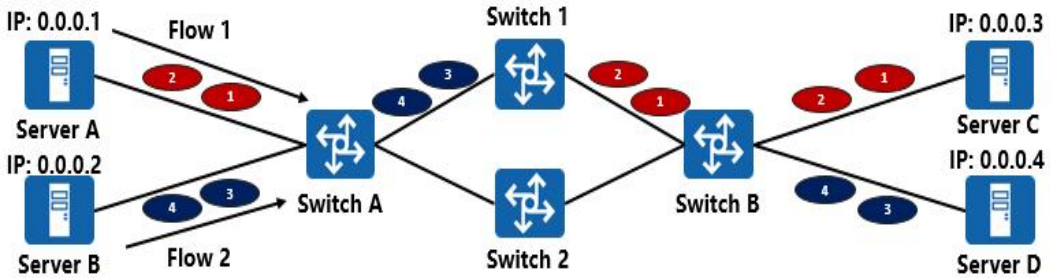
Technologies to meet AI fabric requirements must consider following communication characteristics.

- Periodic communication is consistent in each iteration with burst traffic pattern
- Small number of traffic flows with large size for a single flow
- GE~100s GE level communication for each iteration

Load Balancing Challenges in AI Fabric

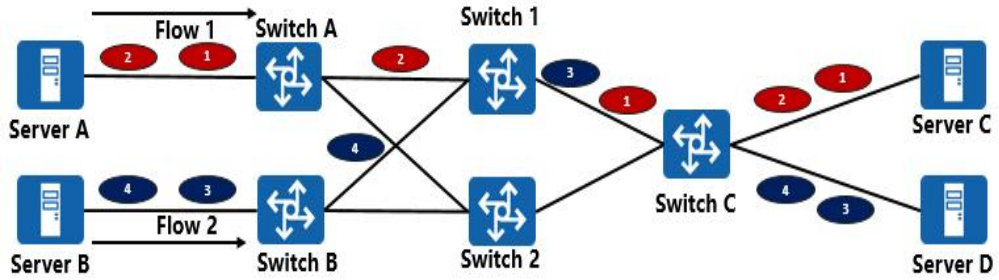
Flow based ECMP is poor of handling asymmetric AI traffic.

Imbalanced load traffic leads to congestion and low throughput.



Local collision:

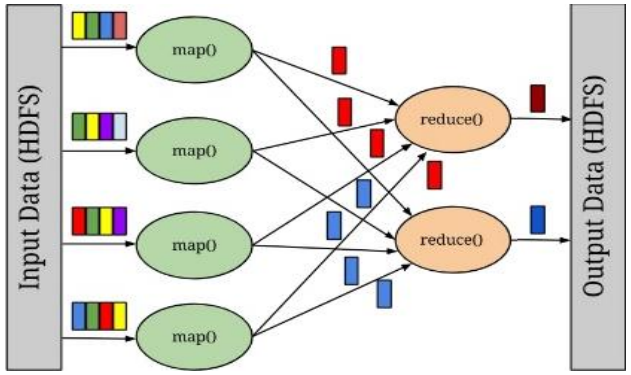
5 tuple based hash algorithm may output the same hash-key for different flows, resulting multiple flows to be forwarded to the same path causing local collision.

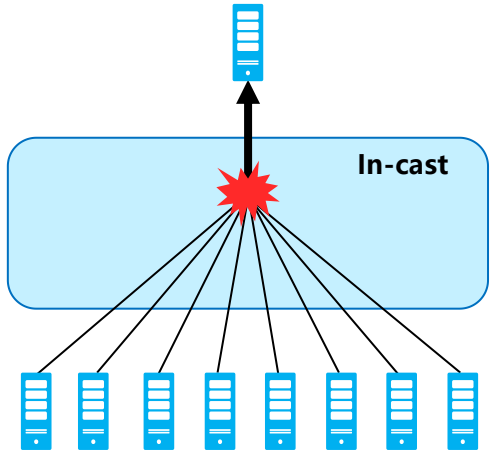


Downstream collision:

The local decision-making mechanism lacks of global view of the fabric (e.g. downstream nodes status) which may select multiple flows forwarded to the same downstream path, causing downstream collision.

FC/CC Challenges in AI Fabric





In-cast traffic (‘reduce’ operation) exists in AI fabric and can easily create congestion.

Current solution

Sender push traffic into network until notification of congestion is received , then sender adjust sending rate

- End-to-end congestion control (e.g. DCQCN)
- Hop-by-hop flow control(e.g. PFC)

Major issue

Passive control of congested flows may cause GPU to idle, thus waste GPU resource

Summary

- **Introduce communication pattern in large-scale AI model**
- **Analyze AI fabric requirements in order to shorten communication time.**
  - Notes: Besides communication time, AI fabric has additional requirements in other aspects, such as reliability, security and maintainability. Those are not covered in this presentation.
- **List key technologies to meet AI fabric requirements**
- **Focus on current LB/CC/FC challenges when considering AI communication characteristics**

Thank You !