

Reviewing Load balancing issues in AI Computing Network

Jieyu Li (China Mobile)

November Plenary 2024

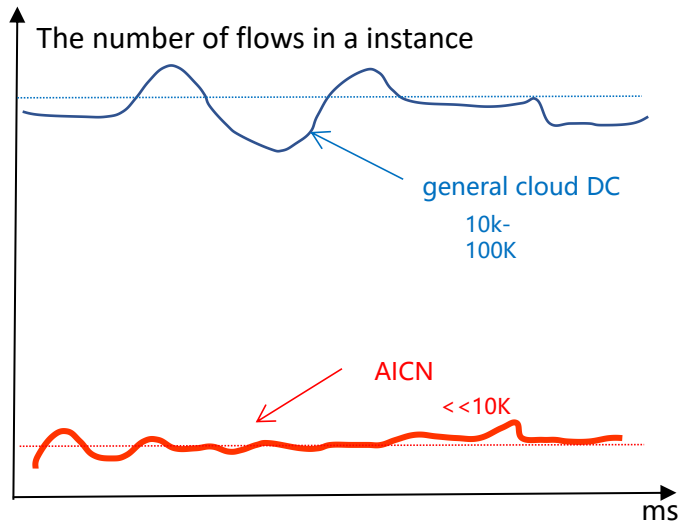
Purpose

- This contribution is related to the topic of load balancing in AICN study item.
- Reviewing the load balancing issues in AICN.

Traffic Pattern of AI Computing Network

- With the wide deploying of LLM, the traffic pattern in AI network is clear

Low entropy



- **General Cloud DC:** Great fluctuation, 10K~200K^[1]
- **AICN:** Relatively stable, a few dozen to hundreds of connections^[1]

Elephant flows

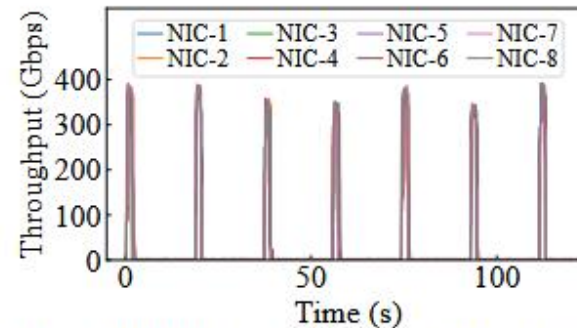
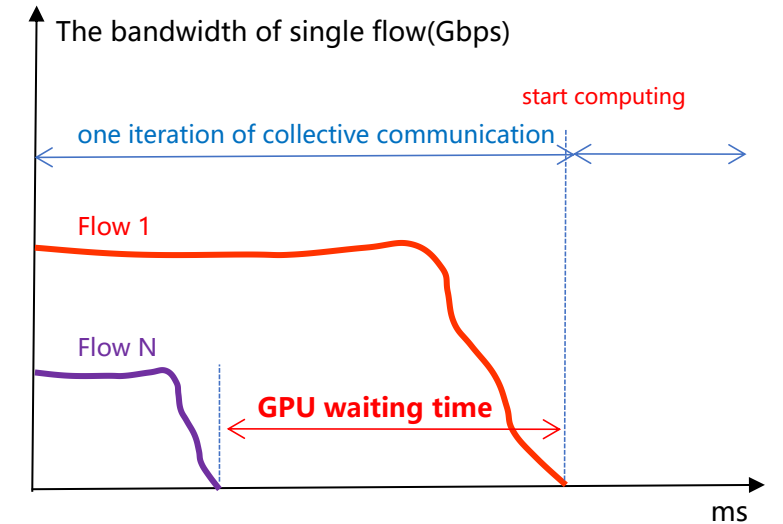


Figure 2: NIC egress traffic pattern during production model training.

from Sigcomm2024: Alibaba HPN: A Data Center Network for Large Language Model Training

- **General Cloud DC:** Low bandwidth flows^[3]
- **AICN:** High bandwidth flows^[3], periodic burst

Collective communication



- **General Cloud DC:** Point-to-Point dominant
- **AICN:** Collective communication dominant

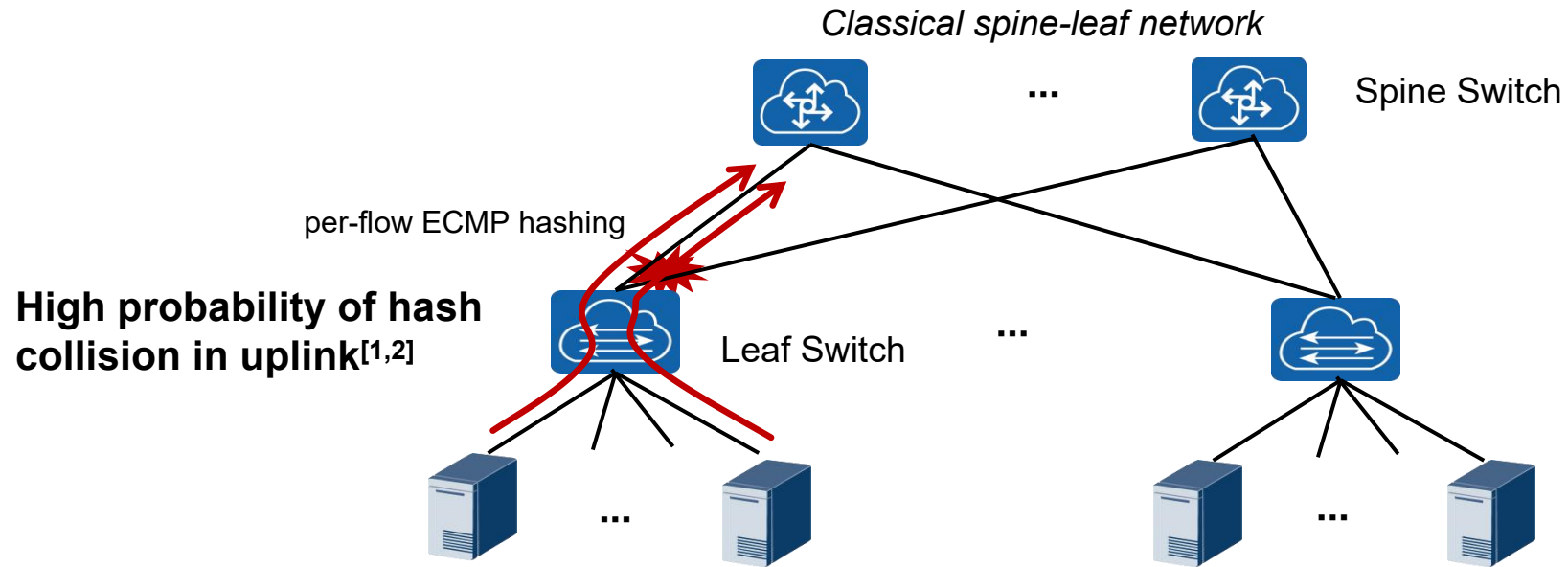
[1] Qian K, Xi Y, Cao J, et al. Alibaba hpn: A data center network for large language model training[C]//Proceedings of the ACM SIGCOMM 2024 Conference. 2024: 691-706.

[2] Gangidi A, Miao R, Zheng S, et al. Rdma over ethernet for distributed training at meta scale[C]//Proceedings of the ACM SIGCOMM 2024 Conference. 2024: 57-70.

[3] Nvidia: Next-Generation Networking for the Next Wave of AI.

Load Imbalance Problem in AICN

- **Hash collision** is the main problem when conventional per-flow Equal Cost Multi-Path (ECMP) hashing applying in AICN^[1,2]



- Due to the features of hash algorithm, **low entropy** will lead to the high probability of flows collision.
- Especially for **elephant flows**, the collision will cause more severer imbalance which magnifies long tail latency.
- As AI computing is mainly based on **collective communication**, shorten long tail latency is critical to improve computing efficiency^[3]

[1] Qian K, Xi Y, Cao J, et al. Alibaba hpn: A data center network for large language model training[C]//Proceedings of the ACM SIGCOMM 2024 Conference. 2024: 691-706.

[2] Gangidi A, Miao R, Zheng S, et al. Rdma over ethernet for distributed training at meta scale[C]//Proceedings of the ACM SIGCOMM 2024 Conference. 2024: 57-70.

[3] Cisco:Evolve your AI/ML Network with Cisco Silicon One.

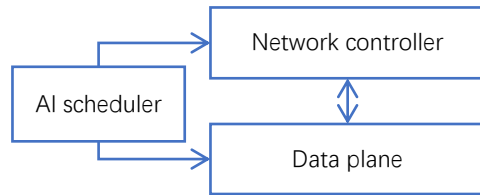
Existing Improvement Options

- There several solutions to release the load balancing problem in AICN, here sort out the existing solutions based on public materials or papers:

Per-Flow-based Solutions

1. Path planning: Set a definite path for each flow, avoiding uncertain hash collision.

- e.g., centralized traffic engineering.^{[1][2]}



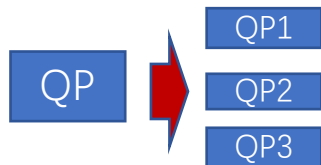
Limitations:

- “lower performance when multiple links failures happen”^[1]
- “additional software complexity and manageability overhead.”^[1]
- hard to scale.

...

2. Splitting into subflows: Make an elephant flow into multiple subflows to enlarge entropy^[2], lowering the possibility of hash collision.

- e.g, QP scaling^[1]



Limitations:

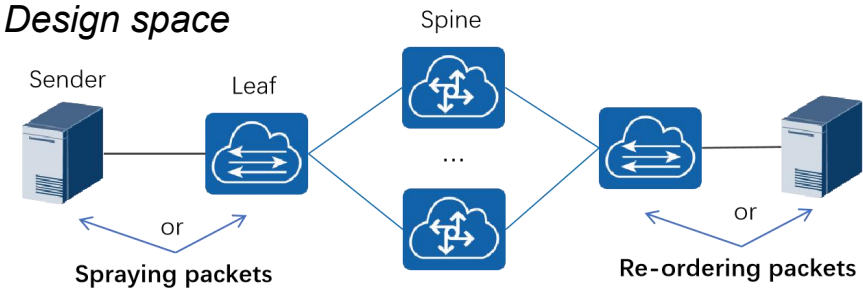
- “the underlying probabilistic nature of hashing was a persistent downside”^[1]

...

Per-Packet-based Solutions

3. Packet spraying: Distribute packets of a flow into multiple paths.

Design space



- a1) Switch-based spraying
- a2) Host-based spraying
- b1) Switch-based re-ordering
- b2) Host-based out-of-order packets processing

- a1-b1: switched-based solution, decoupling with NIC
 - e.g., Cisco’s DSF^[3],...
- a1-b2: cooperation of NIC and switch.
 - e.g., Nvidia’s Spectrum-X+BlueField3^[4],...
- a2-b2: host-based solution with low requirements for switches.
 - e.g., some new transportation protocols...

[1] Gangidi A, Miao R, Zheng S, et al. Rdma over ethernet for distributed training at meta scale[C]//Proceedings of the ACM SIGCOMM 2024 Conference. 2024: 57-70.

[2] Huawei: <https://info.support.huawei.com/info-finder/encyclopedia/en/NSLB.html>

[3] Cisco: <https://www.ciscolive.com/c/dam/r/ciscolive/global-event/docs/2024/pdf/AIHUB-1004.pdf>

[4] Nvidia: Next-Generation Networking for the Next Wave of AI.

Conclusion

- Load balancing is still one of the central issues in AICN.
- There are lots of load balancing solutions for AICN that can be classified into flow-based and packet-based, and per-packet load balancing is the direction that attracts more attention.
- It's necessary to give a deep analysis on per-packet-based solutions' pros and cons, application scenarios in AICN study item report.

Thank You !