APNet 2025
SHANG HAI

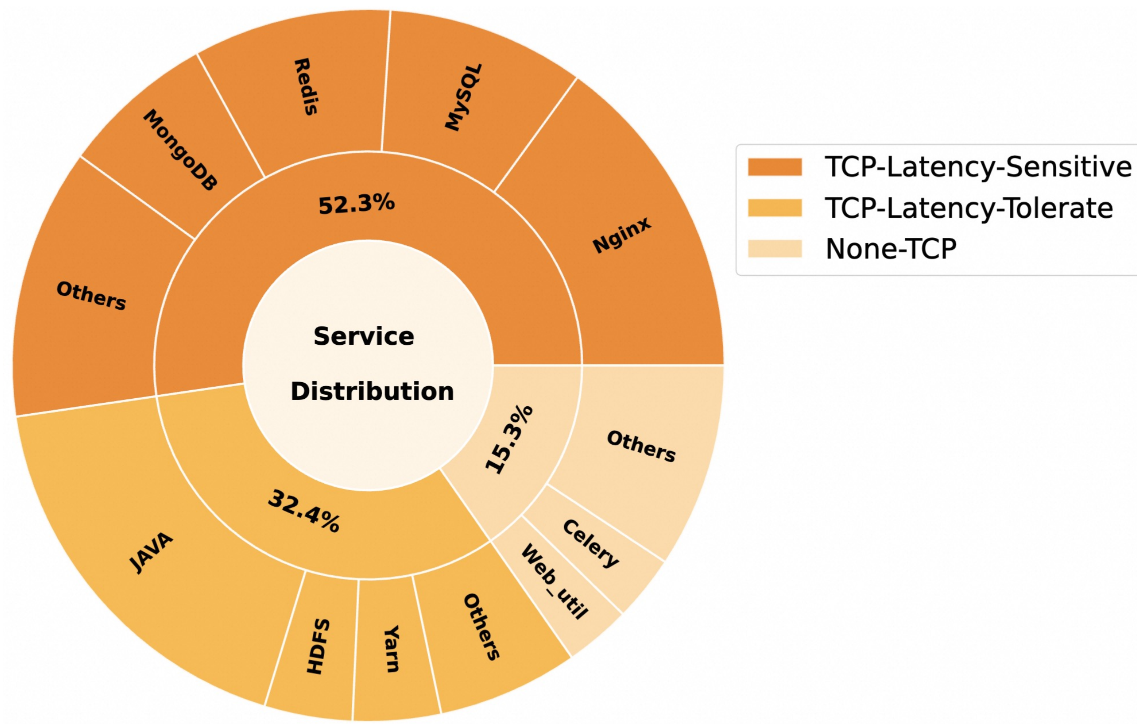
Understanding the Long Tail Latency of TCP in Large-Scale Cloud Networks

Zihao Fan, Enge Song, Bo Jiang, Yang Song, Yuke Hong, Bowen Yang, Yilong Lv,
Yinian Zhou, Junnan Cai, Chao Wang, Yi Wang, Yehao Feng, Dian Fan, Ye Yang, Shize Zhang,
Xiaoqing Sun, Jianyuan Lu, Xing Li, Jun Liang, Biao Lyu, Zhigang Zong, Shunmin Zhu

Shanghai Jiao Tong University · Alibaba Cloud · Zhejiang University

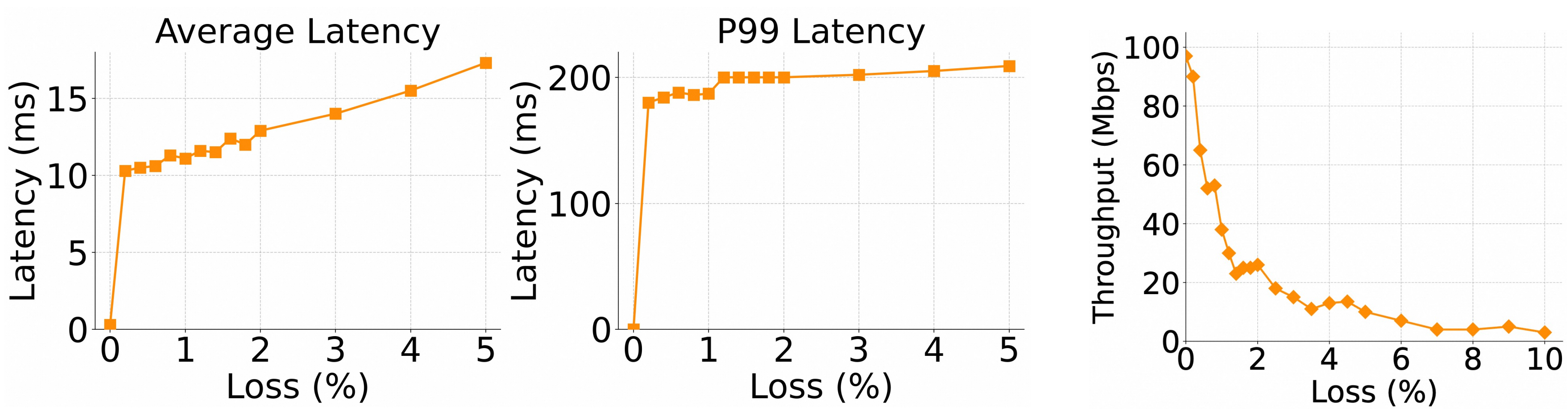
Why TCP's Tail Latency Matters in the Cloud

- **TCP** dominates ECS:
  - **85%+** of workloads rely on it
- **60%** of TCP services are **tail-sensitive**
  - (e.g., Redis)
- Long tail latency
  - → SLA violation
  - → Revenue loss



Why TCP's Tail Latency Gets Worse in Cloud

- **Single-path** transmission can overload paths
- **Loss detection** slow (RTO ~200ms, triple dup ACK)
- **Congestion control** reacts late (loss-driven)



Empirical: 1% loss → P99 latency increases 150×

Why is TCP's Tail Latency So Hard to Eliminate

Cloud-scale:

O(1M) of links,

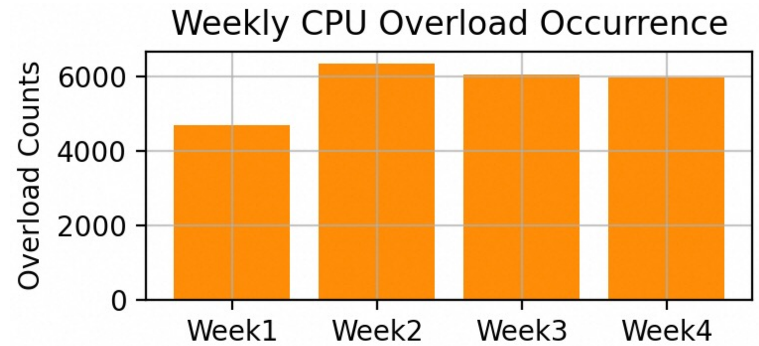
O(100k) servers per region

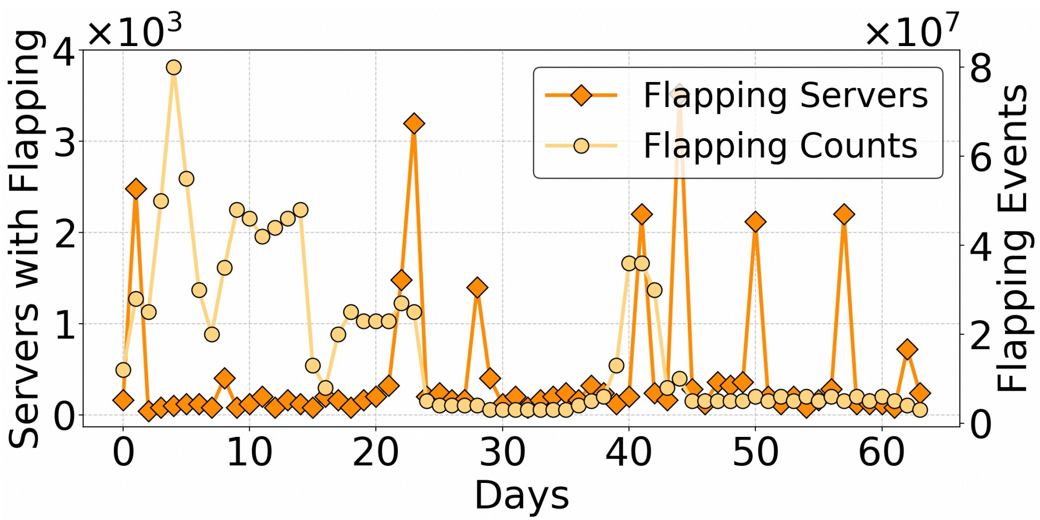
→

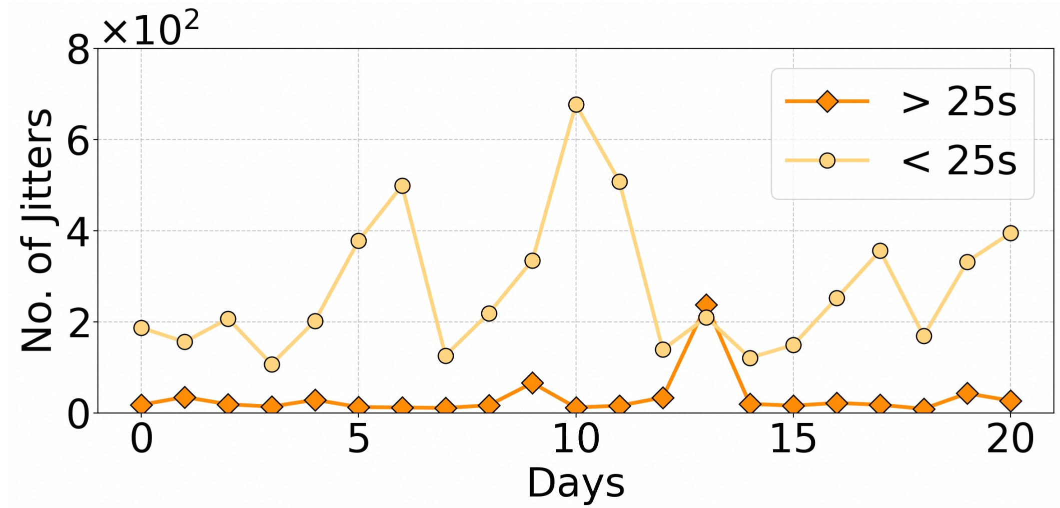
Network Instabilities are common and **inevitable**

Root Causes: Network Instabilities

- Elephant flows: **thousands** of times/week
- NIC flapping: **millions** of times/day
- Network jitter: **hundreds** of times/day

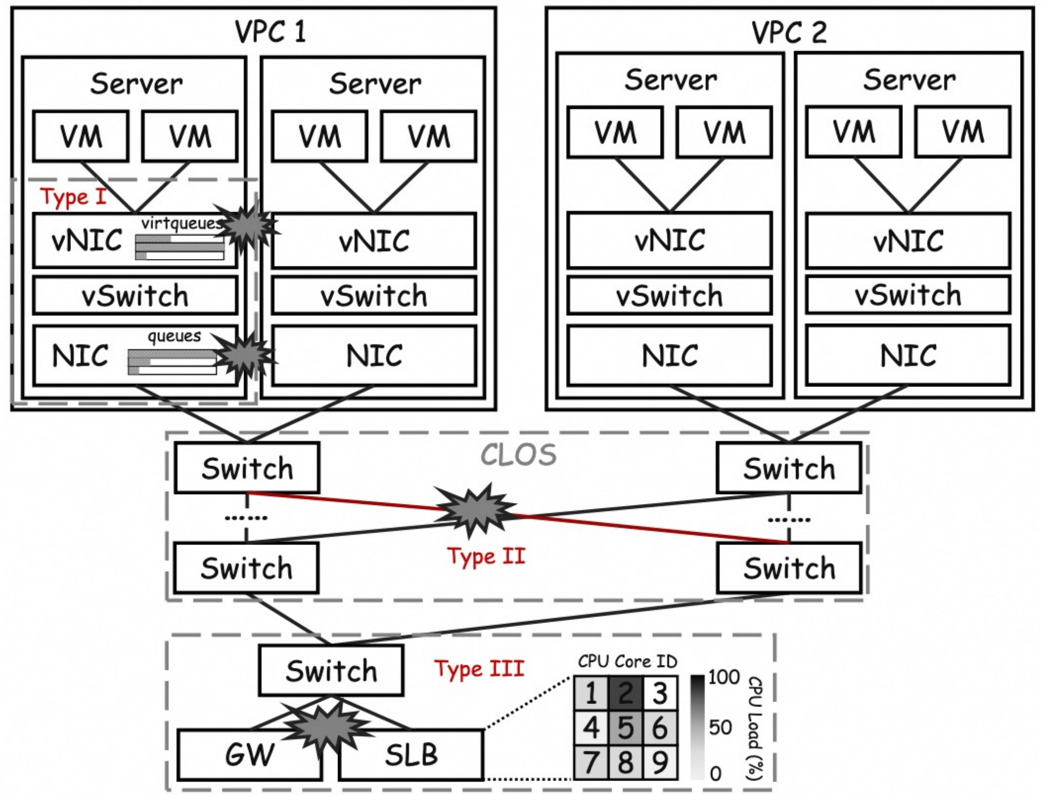






Result of Instability: Packet Loss

- Packet loss occurs at **multiple points**
  - Type I: Packet loss in a single server
    - Physical NIC
    - Front-end (vNIC) and back-end (vSwitch)
  - Type II: Packet loss in physical networks
    - Physical Link
    - Switch
  - Type III: Packet loss in Middlebox
    - Gateway, Load balancer, NAT, …

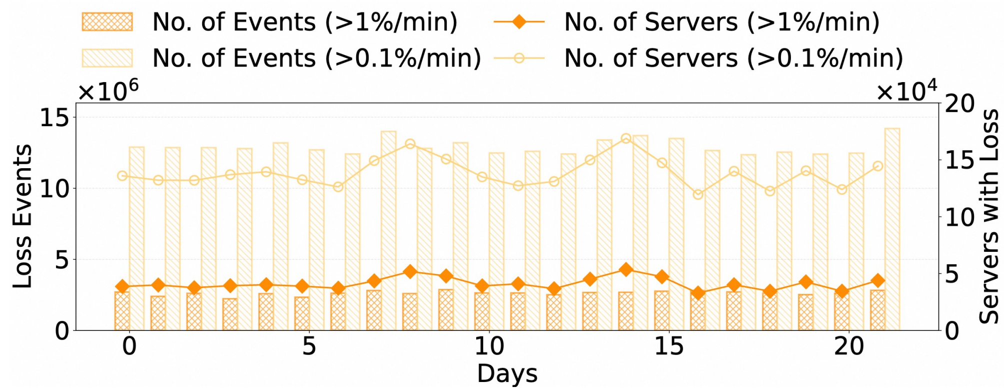


Result of instability: Packet Loss

Type I: Packet loss in a single server

Type III: Packet loss in Middlebox

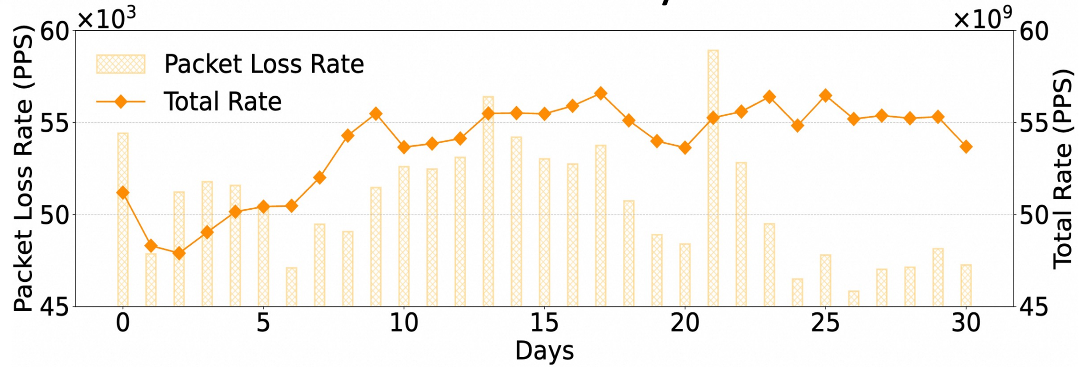
Physical NIC



vNIC front-end and vSwitch back-end

Solution	*Average*	*P90*	*P99*	*P999*
Kernel-based VM	0.015%	0%	0.17%	2.0%
SmartNIC I	0.006%	0%	0.02%	1.0%
SmartNIC II	0.072%	0.03%	0.29%	2.0%

Cloud Gateway



Why Existing Solutions Fall Short

Our Goal

Mitigating long tail latency in **unstable, large-scale cloud networks** while maintaining **complete transparency to end users**

Limitations

Limited Performance Improvement
- Coarse-grained multipath[1] [2]
- Lack of receiver-side reordering[3][4]
- Random path selection[1][5]

Intrusiveness to Users
- Dependency on ECN[6][7]
- Kernel modifications at end hosts[1][8]

Poor Compatibility and Scalability
- Custom switch functionalities[9][10]
- Centralized control plane[11][12]

[1] Qureshi et al. PLB: Congestion signals are simple and effective for network load balancing. SIGCOMM 2022, pp. 207–218.
[2] Google Cloud. *Introducing Falcon: A reliable, low-latency hardware transport.* Google Cloud Blog, 2023.
[3] Shalev et al. *The Tail at AWS Scale.* IEEE Micro, 2024.
[4] Le et al. *STrack: A Reliable Multipath Transport for AI/ML Clusters.* arXiv:2407.15266, 2024.
[5] Vanini et al. *Let it flow: Resilient asymmetric load balancing with flowlet switching.* NSDI 2017, pp. 407–420.
[6] Katta et al. *Clove: Congestion-aware load balancing at the virtual edge.* CoNEXT 2017, pp. 323–335.
[7] Kabbani et al. *Flowbender: Flow-level adaptive routing for improved latency and throughput in datacenter networks.* CoNEXT 2014, pp. 149–160.
[8] Ford et al. *TCP extensions for multipath operation with multiple addresses.* Technical Report, 2013.
[9] Alizadeh et al. *CONGA: Distributed congestion-aware load balancing for datacenters.* SIGCOMM 2014, pp. 503–514.
[10] Song et al. *Network Load Balancing with In-network Reordering Support for RDMA.* SIGCOMM 2023, pp. 816–831.
[11] Al-Fares et al. *Hedera: Dynamic flow scheduling for data center networks.* NSDI 2010, pp. 89–92.
[12] Curtis et al. *Mahout: Low-overhead datacenter traffic management using end-host-based elephant detection.* IEEE INFOCOM 2011, pp. 1629–1637.

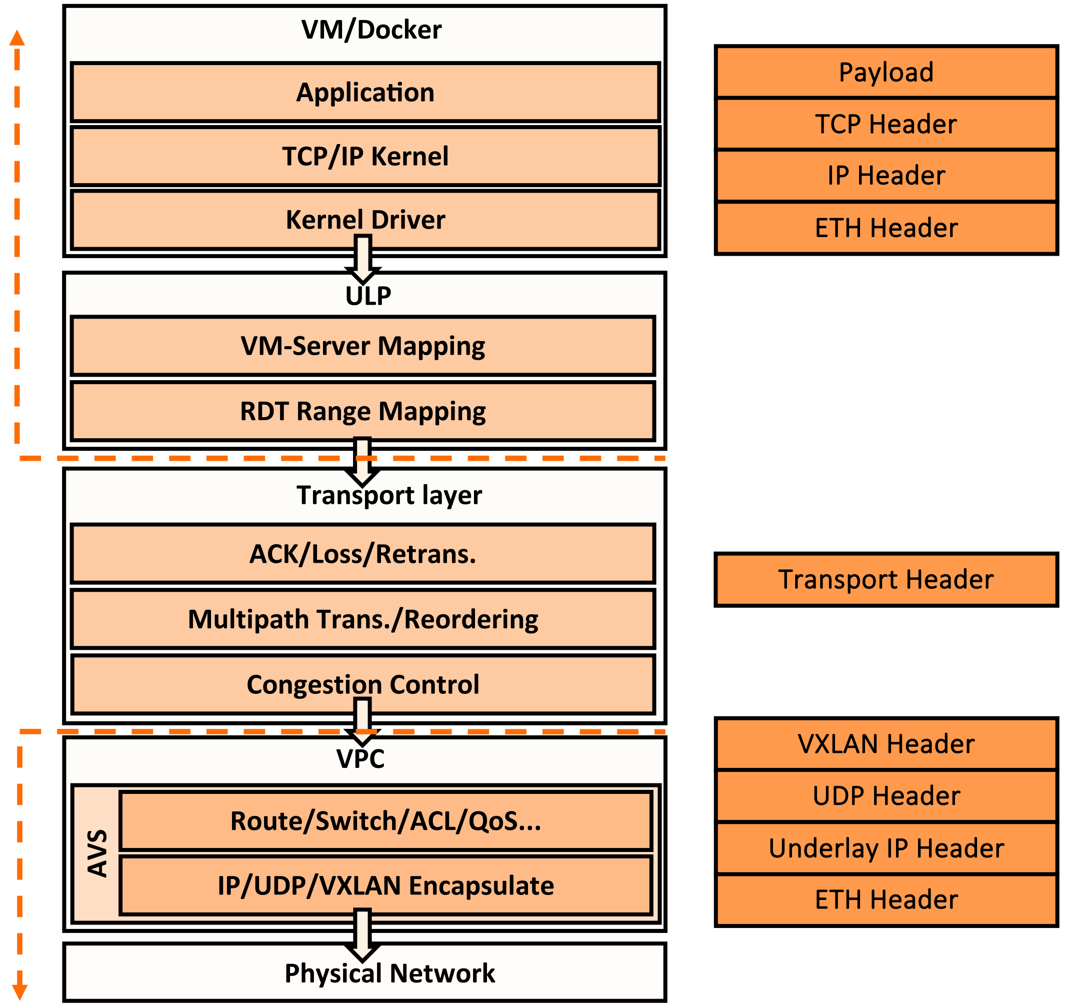
Bifrost

- RTT-Aware multipath transmission
- Hybrid hardware-software reordering
- ACK aggregation via delayed bitmap

1. High performance guarantee

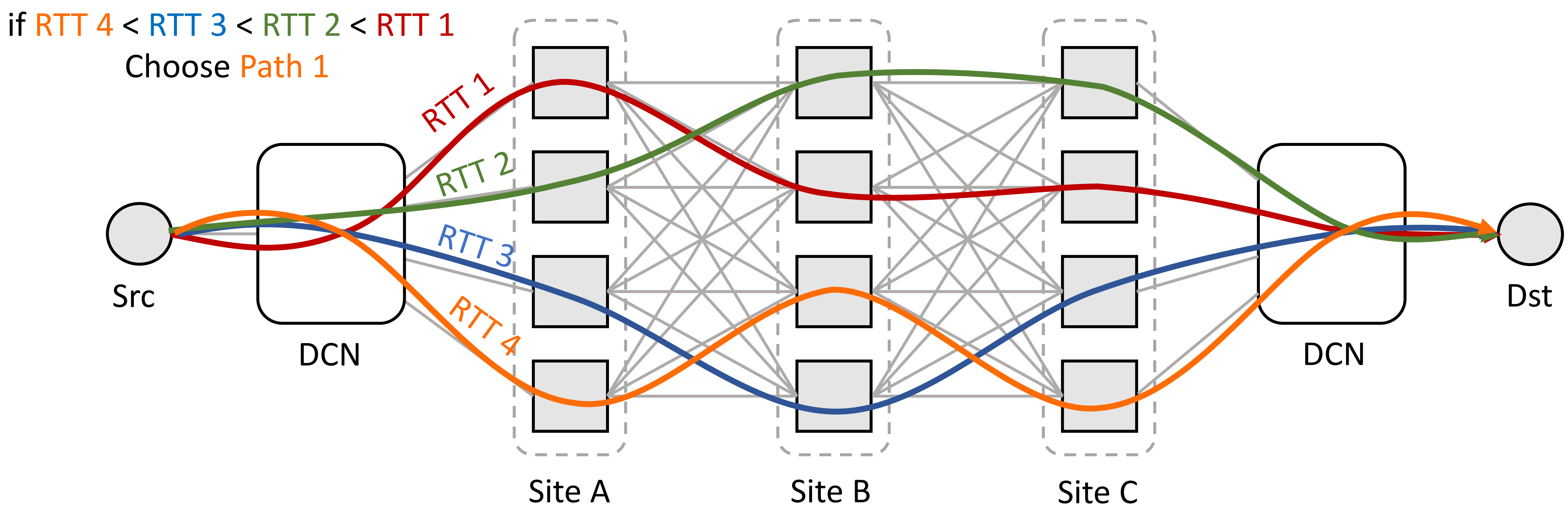
2. Non-intrusive to users

3. W/o requiring support from network devices



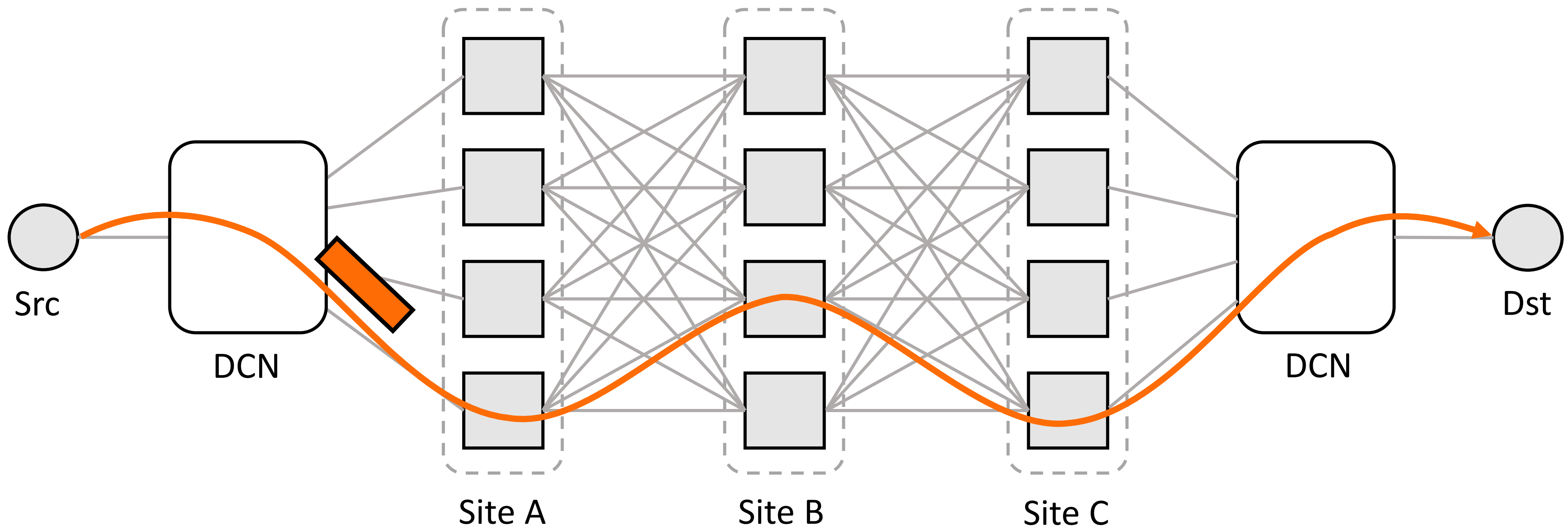
RTT-Aware Multipath Transmission

- Partition flow into equal-sized packet groups for scheduling
- Dynamically select the **lowest RTT paths** across groups



RTT-Aware Multipath Transmission

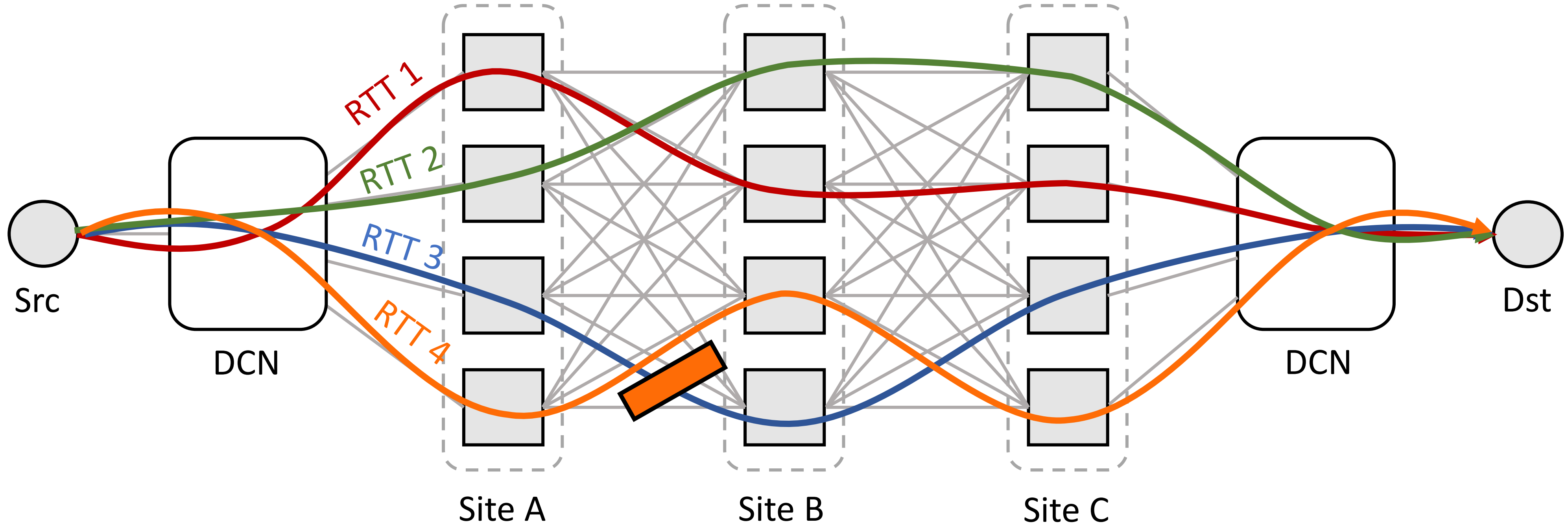
- Partition flow into equal-sized packet groups for scheduling
- Dynamically select the **lowest RTT paths** across groups



RTT-Aware Multipath Transmission

- Path stickiness optimization

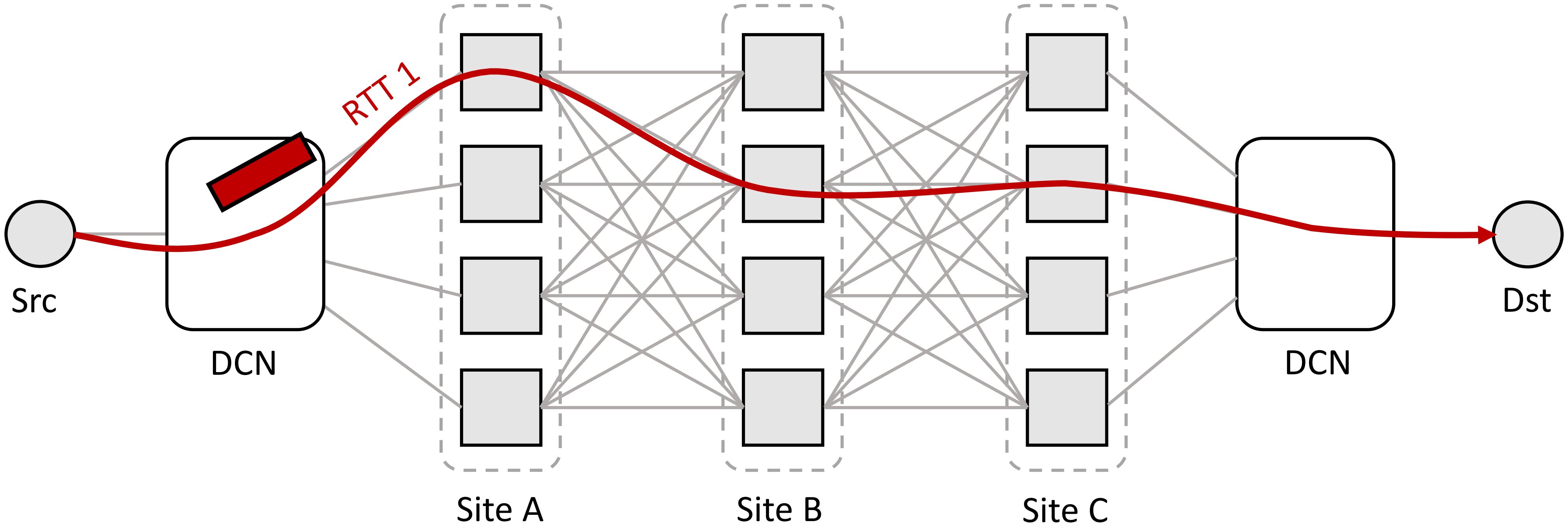
min(RTT 3,RTT 2,RTT 1) = RTT 1
If RTT 4 - RTT 1 > δ
Switch to Path 1



RTT-Aware Multipath Transmission

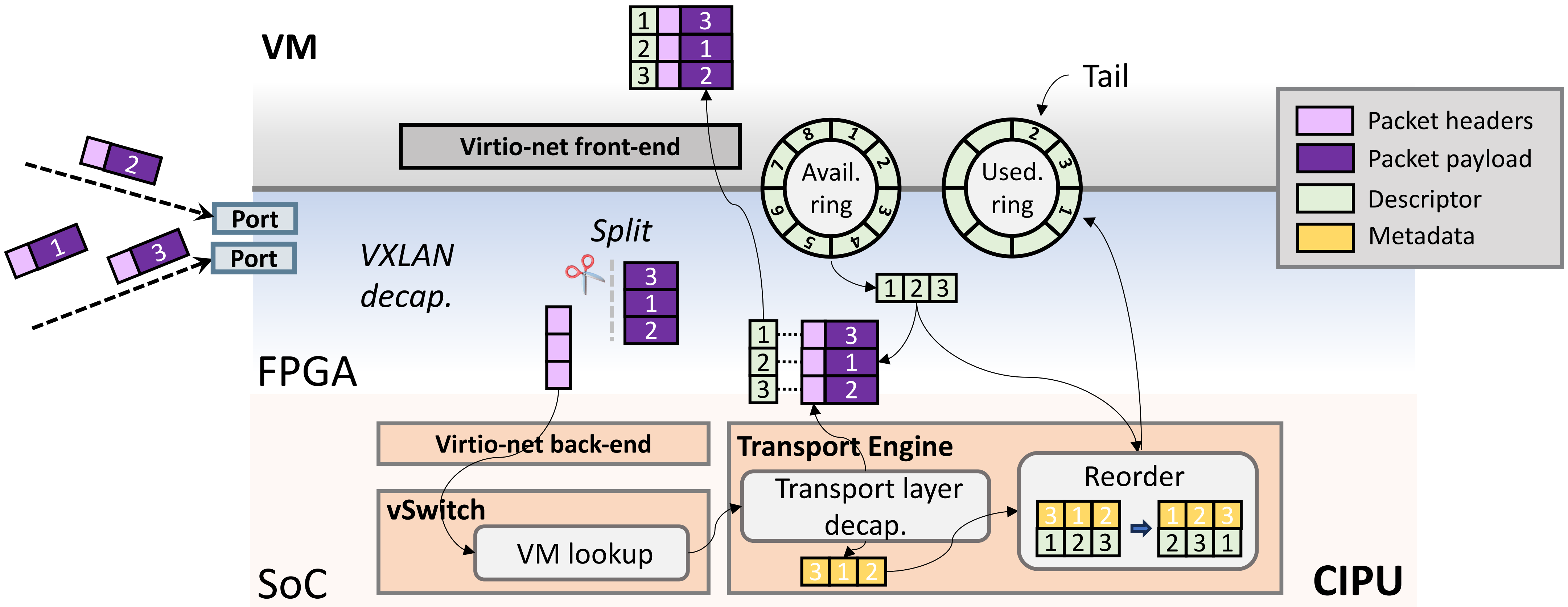
- Path stickiness optimization

min(RTT 3,RTT 2,RTT 1) = RTT 1
If RTT 4 - RTT 1 > δ
Switch to Path 1



Hybrid Hardware-Software Reordering

- Hardware receives packets
- Software enforces reordering via metadata.

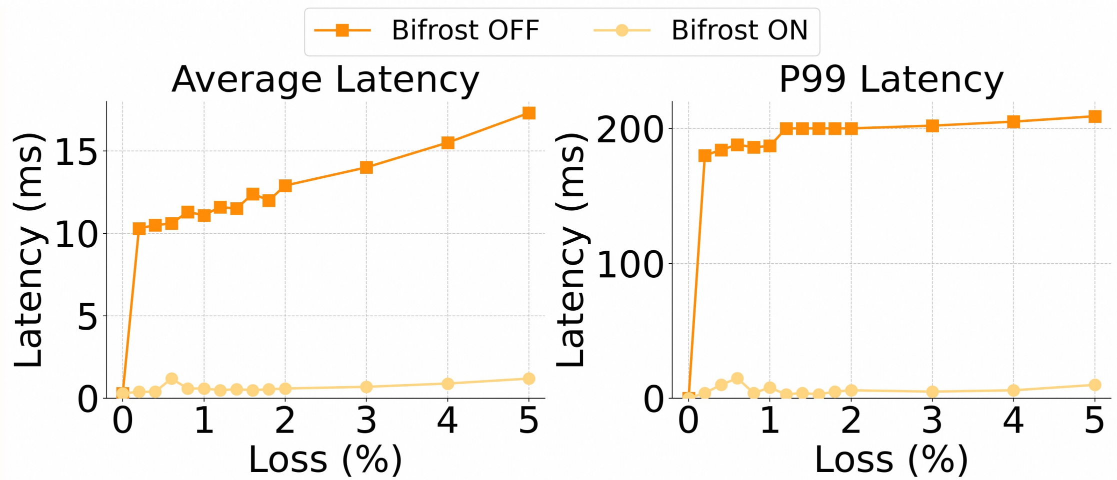


ACK Aggregation Via Delayed Bitmap

- **Delayed ACK generation** to capture vNIC losses
  - ACK is sent **only after packet reaches the VM**, not at physical NIC reception
- ACK format follows **bitmap ACK standard**
  - Sender triggers fast retrans. upon detecting **gaps in the bitmap**
  - ACK(3|0100) -> retrans. packet 3
- Precise Retransmission Timeout (RTO)
  - RTO is set to **4ms**, roughly 2× RTT in data center networks
  - Significantly faster than traditional TCP RTO (200ms)

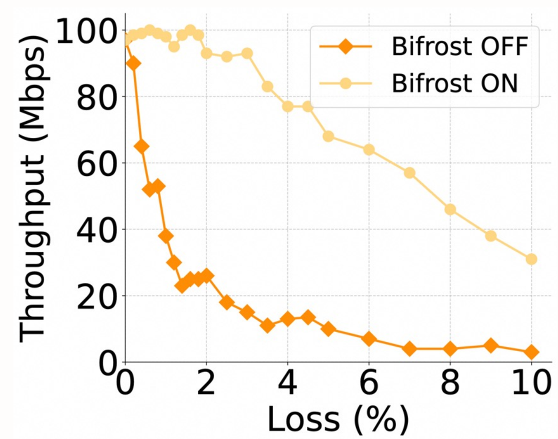
Preliminary Evaluation

Nginx Short Connection



Reduces tail latency by over **90%** under varying packet loss

Sockperf



Even under 1% packet loss,
Bifrost sustains **97%** of peak throughput

Conclusion

- **Cloud-scale network instability causes frequent tail latency spikes in TCP services**, impacting SLA-critical applications like Redis.
- **We present Bifrost, a scalable and non-intrusive transport layer** that combines RTT-aware multipath, hybrid reordering and delay ACK.
- **Evaluation shows Bifrost reduces P99 latency by >90% and sustains 97% throughput under loss**, significantly improving both transport and application performance.

Understanding the Long Tail Latency of TCP in Large-Scale Cloud Networks

Q & A