The 2025 USENIX Annual Technical Conference (ATC 2025)

FLB: Fine-grained Load Balancing for Lossless Datacenter Networks

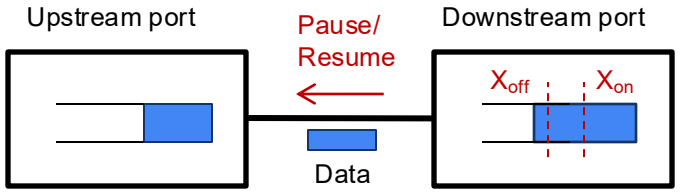
Jinbin Hu[1,2,3*] , **Wenxue Li**[2*] , Xiangzhou Liu[2] , Junfeng Wang[2] , Bowen Liu[2] , Ping Yin[4] , Jianxin Wang[1], Jiawei Huang[1], Kai Chen[2]

[1]Central South University
[2]iSING Lab, Hong Kong University of Science and Technology
[3]Changsha University of Science and Technology
[4]Inspur

*Equal contribution

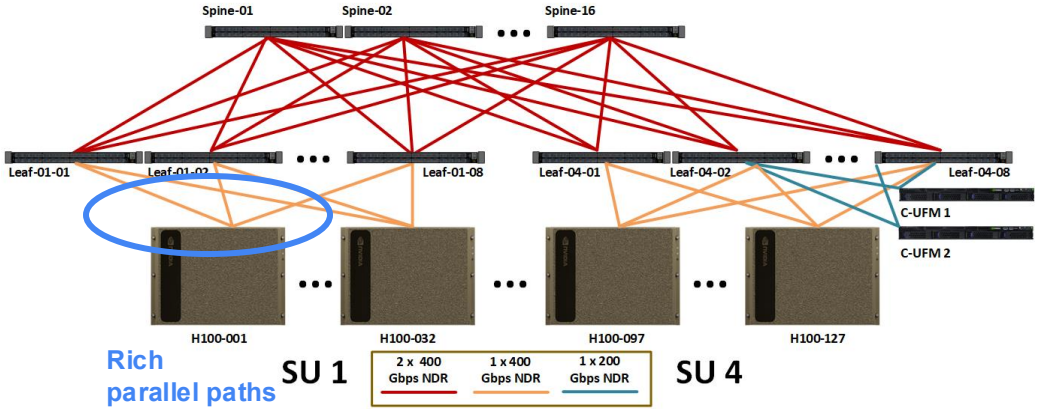
Lossless RDMA Network and Load Balancing

- RDMA over Converged Ethernet (RoCE) is widely deployed
  - Microsoft Azure [1]; Alibaba Cloud [2]; Google cloud [3]
- RoCE employs priority flow control (PFC) to enable a lossless fabric



PFC illustration

- Load balancing (i.e., multipath transmission) is important because modern datacenters usually provide multiple parallel paths



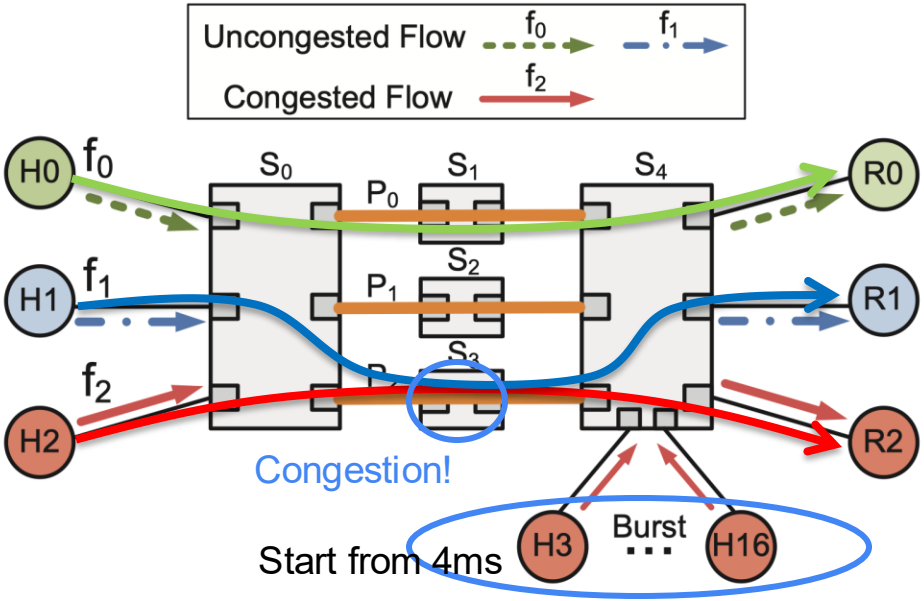
Nvidia Superpod [4]

[1] https://azure.microsoft.com/en-us/blog/azure-linux-rdma-hpc-available
[2] https://www.alibabacloud.com/product/scc
[3] https://cloud.google.com

[4] https://docs.nvidia.com/dgx-superpod/reference-architecture-scalable-infrastructure-h100/latest/network-fabrics.html

Existing LB Schemes are Inefficient in Lossless RDMA Networks – Reason #1

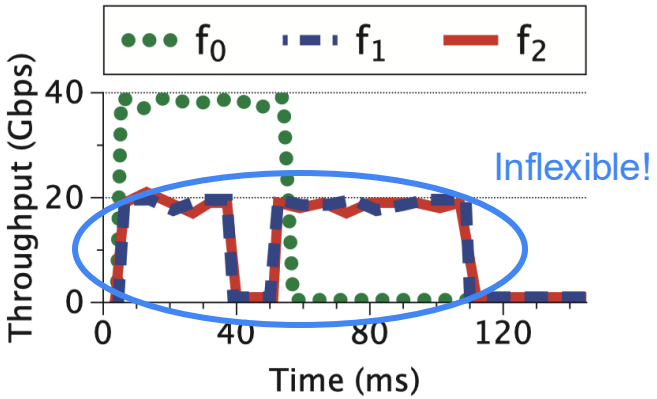
- **Inflexible rerouting leads to load imbalance and link under-utilization**



Small-scale Testbed Experiments.
Transport implemented by using DPDK and P4 switch.

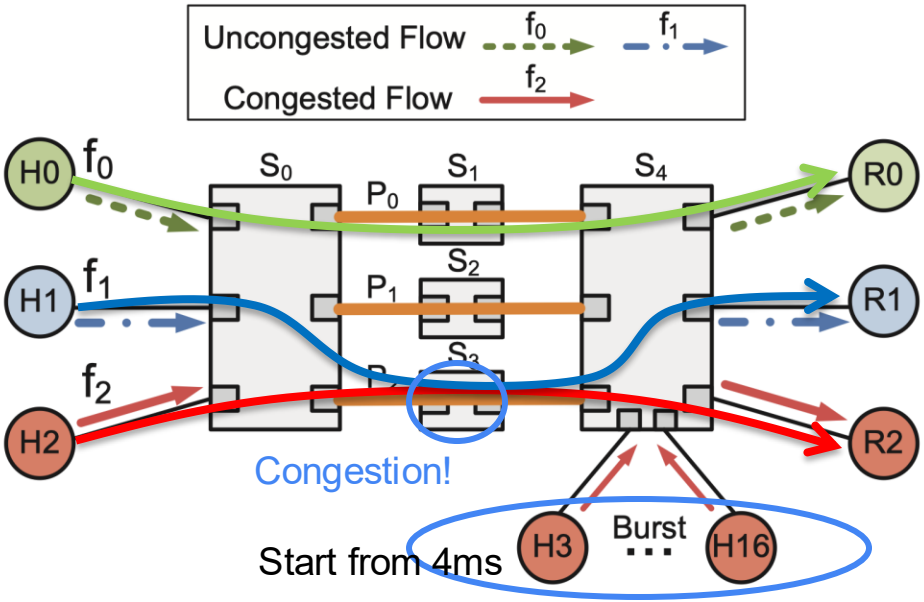
ECMP:

- In some round, *f1* and *f2* are hashed to P2 by coincidence, causing congestion.
- ECMP cannot reroute *f1/f2* after congestion



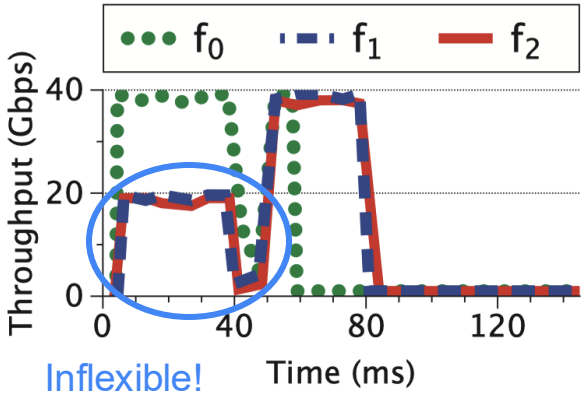
Existing LB Schemes are Inefficient in Lossless RDMA Networks – Reason #1 (Cont.)

- **Inflexible rerouting leads to load imbalance and link under-utilization**



LetFlow[1] (Flowlet-based LB):

- The entire *f1/f2* is treated as one flowlet.
- In some round, f1 and f2 are mapped to P2 by coincidence, causing congestion
- LetFlow cannot reroute *f1* after congestion due to the lack of flowlet in RDMA traffic[2].

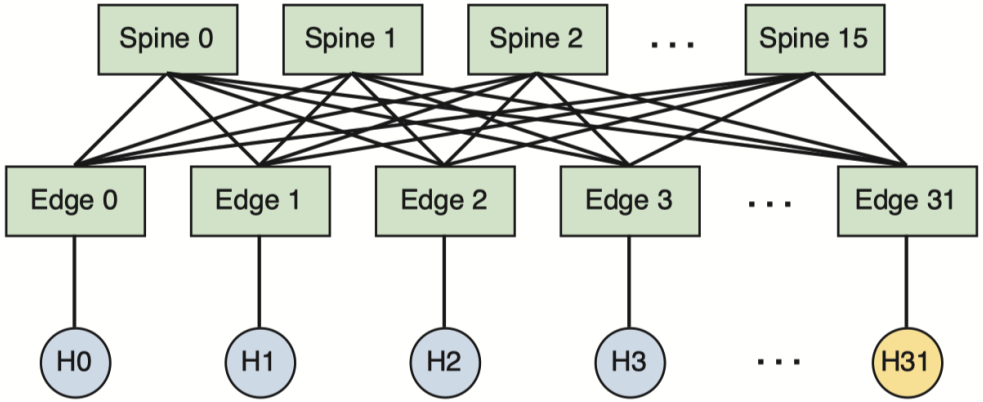


[1] Let it flow: resilient asymmetric load balancing with flowlet switching, NSDI 2017
[2] Multi-Path Transport for RDMA in Datacenters, NSDI 2018

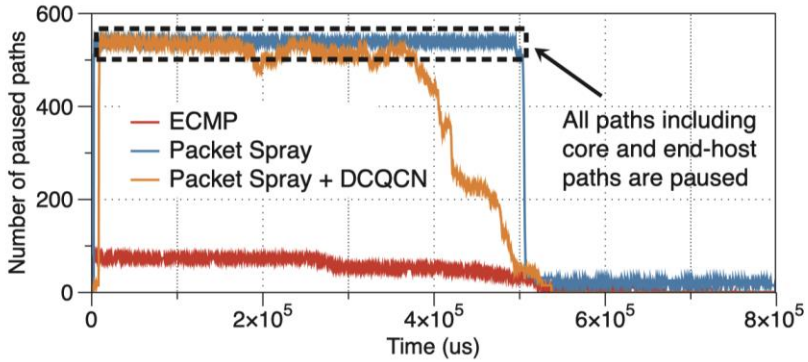
Existing LB Schemes are Inefficient in Lossless RDMA Networks – Reason #2

- **Multi-path transmission expands the influence scope of PFC's HoL blocking**

An extreme case:
H0~H30 simultaneously transmit traffic to H31, creating a long-living 31-to-1 incast pattern.



(a) Topology and workload (H0~H30 send traffic to H31)



- ECMP causes about 70 paths being paused.
- Packet spraying results in all paths (about 340 paths) being paused as the data is spread across all paths
- CC cannot quickly eliminate congestion and stop the PFC pausing

Question:

Can we design a load balancing scheme for PFC-enabled lossless DCNs that achieves high link utilization while eliminating PFC side effects?

Goal #1: Flexibly reroute the traffic to effectively balance load and enhance link utilization

Goal #2: Eliminate the head-of-line (HoL) blocking and congestion spreading during PFC triggering

Goal #3: Reduce dependency on complex congestion control schemes

Key idea of FLB (Fine-grained Load Balancing)

Ideal situation

When there is no congestion:

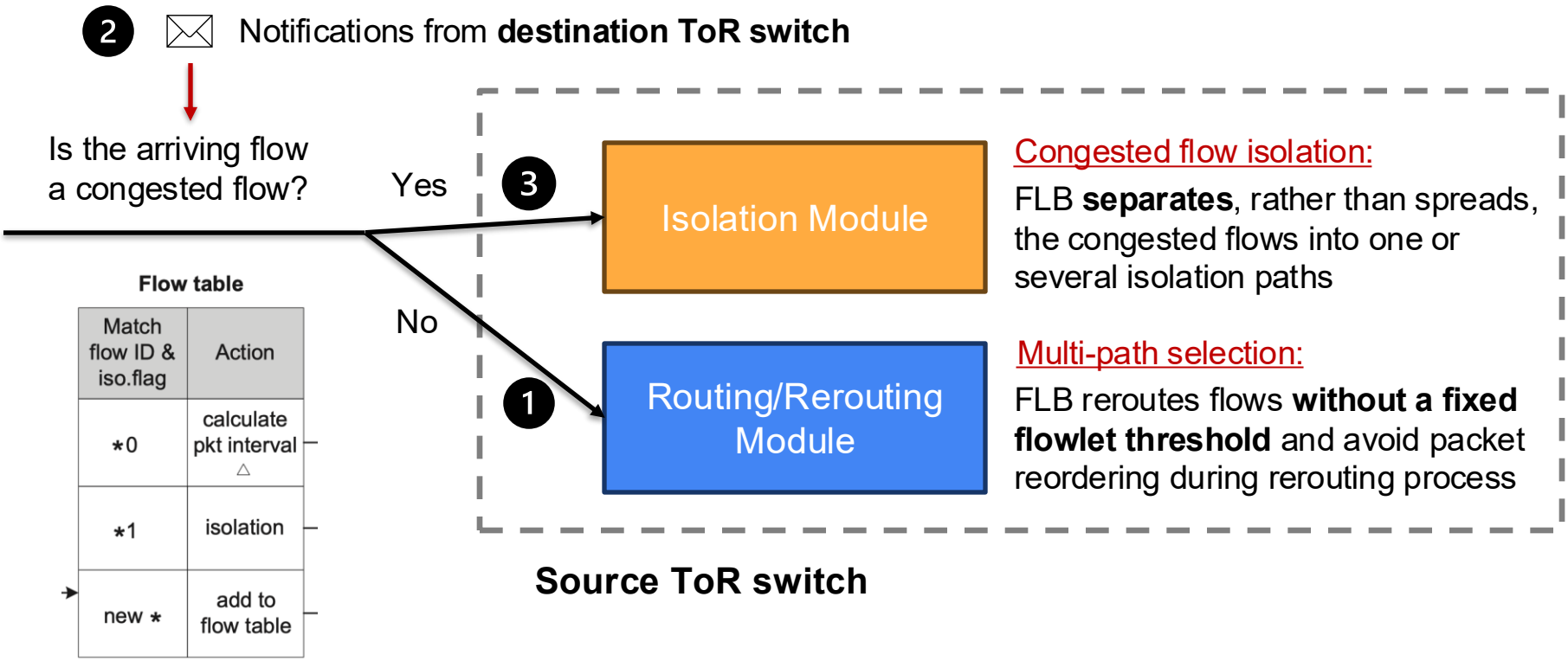
All packets should be transmitted via multiple paths

When congestion happens:

The packets of congested flows should be transmitted using single path to avoid congestion spreading

- ✓ Uncongested flows: multi-path
- ✓ Congested flows: single-path

Key idea of FLB (Fine-grained Load Balancing) (Cont.)



Routing/Rerouting Module

An arriving uncongested flow

Is it a new flow?

No

Select the fastest path with ($P_{owd} < P'_{owd}$) && ($P_{owd} - P'_{owd}$) < ($t'_{cur} - t'_{prev}$)

Yes

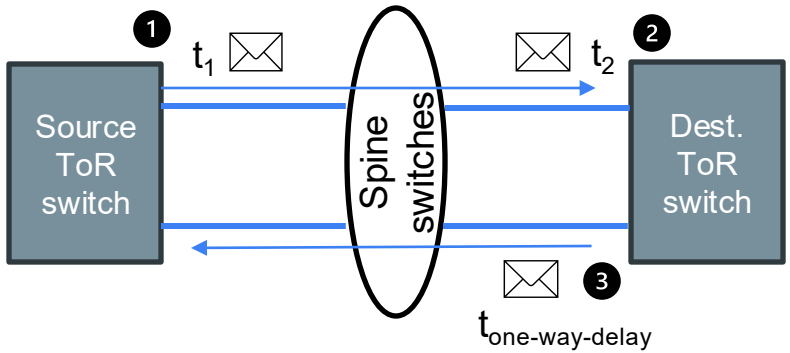
Select a path with the minimum delay for it

P_{owd} and P'_{owd} : Measured one-way delay of the new and current path

t'_{prev} and t'_{cur} : The arrival time of previous and current packet

How to measure <u>one-way delay</u> when the clock is not synchronized?

→ one-way queueing delay

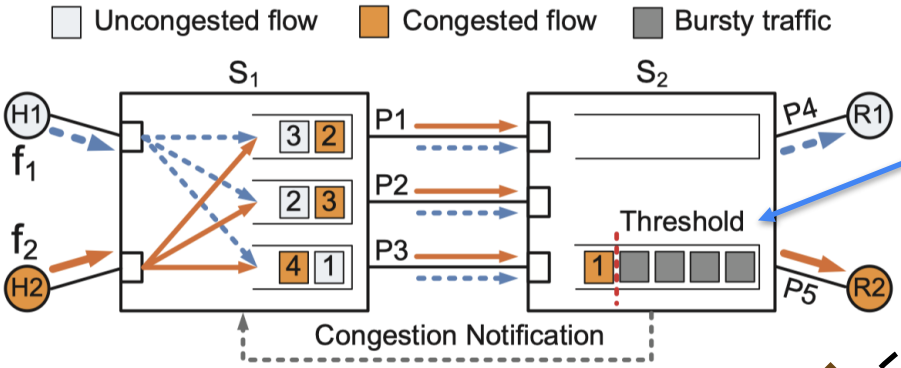


4 $t_{one\text{-}way\text{-}queueing\text{-}delay} = t_{one\text{-}way\text{-}delay} - t_{base\text{-}delay}$

$t_{one\text{-}way\text{-}delay} = t_2 - t_1$

$t_{base\text{-}delay}$ is the one-way delay without any queueing delay (minimum history delay)

Isolation Module



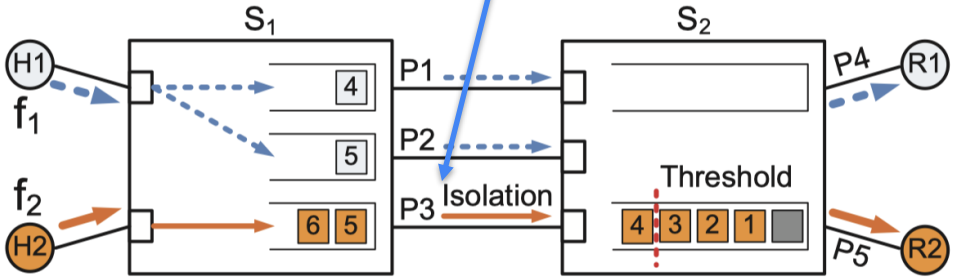
1 When the queue length of egress port reaches the isolation threshold (before PFC is triggered)

Destination ToR Switch

2 A CN frame containing the FID of congestion flows is sent to the source ToR switch

① The source ToR switch stores the FID in a flow table

② The switch will isolate the subsequent congested flow onto a set of isolated paths



Source ToR Switch

Implementation and Evaluation

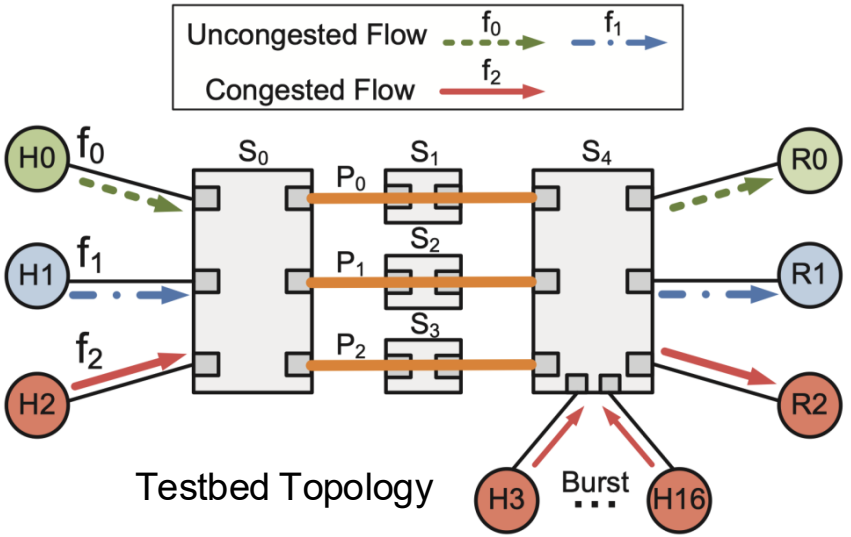
We implement FLB using Wedge 100BF-32X programmable switch

Table 1: Resource consumption of different schemes.

Resource	ECMP	LetFlow	FLB
Match Crossbar	2.41%	4.82%	5.82%
Hash Bits	3.08%	5.67%	5.87%
Gateway	1.39%	2.96%	9.56%
SRAM	1.56%	3.33%	4.12%
VLIW Actions	1.56%	2.34%	3.34%
ALU Instruction	2.6%	5.2%	8.2%

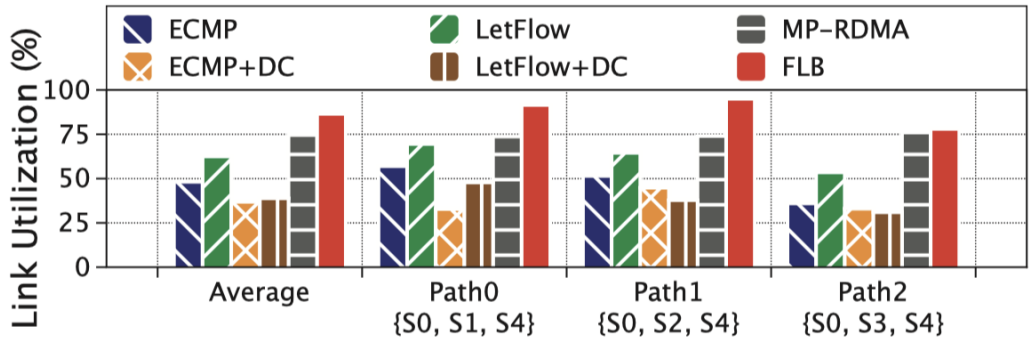
- Testbed server specification:
  - Mellanox ConnectX-5 NICs;
  - DPDK 20.08

- Realistic workload:
  - H0~H16 generate dynamic traffic according to the Web Search workload;
  - H3~H16 are burst flows



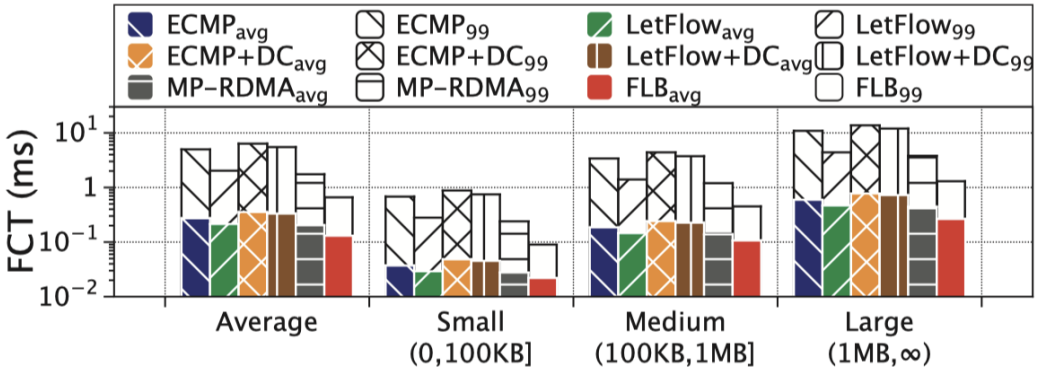
Testbed Topology

Evaluation (Cont.) --- More in paper!



Path level metrics:

FLB achieves higher link utilization on different paths than the other schemes



Flow level metrics:

FLB achieves the lowest average and 99th percentile FCTs of all flows

FLB reduces the AFCT by up to 48%

Thank you!

Please refer to our paper for further information.

Please contact jinbinhu@ust.hk or wlicv@connect.ust.hk if you are interested.