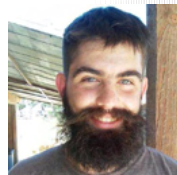
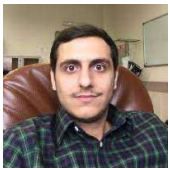
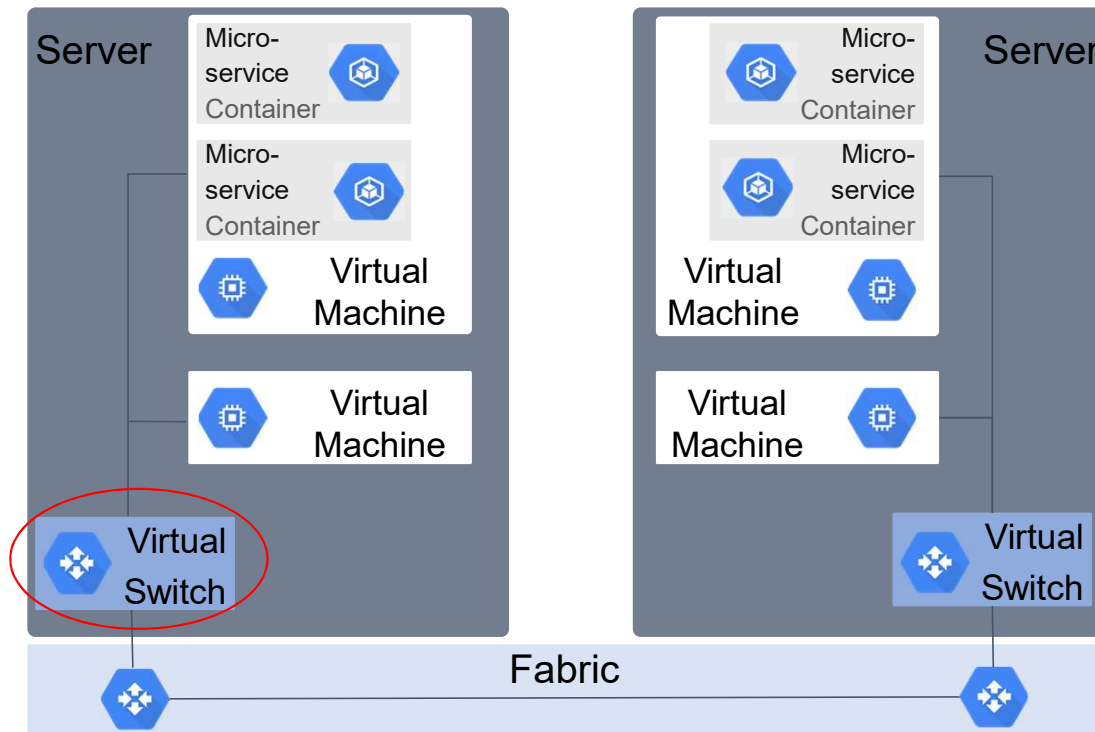


Towards zero-packet loss virtual switches

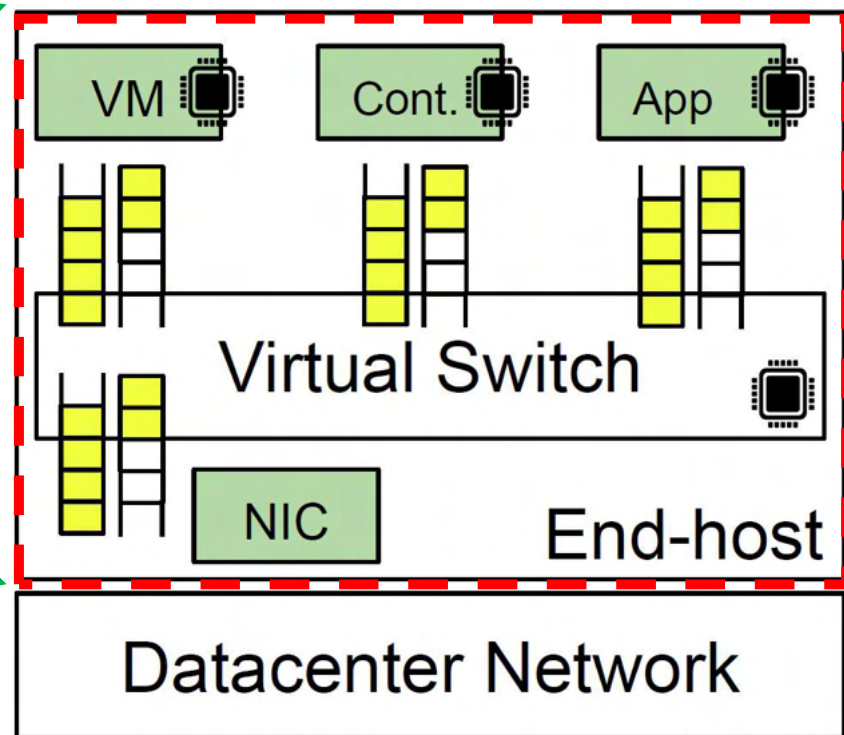
Alireza Sanaee, Farbod Shahinfar, Brent Stephens, Gianni Antichi



Internals of a cloud end-host



A virtual switch



Virtual Switch burns CPU cycles

Virtual switches vs. Physical switches

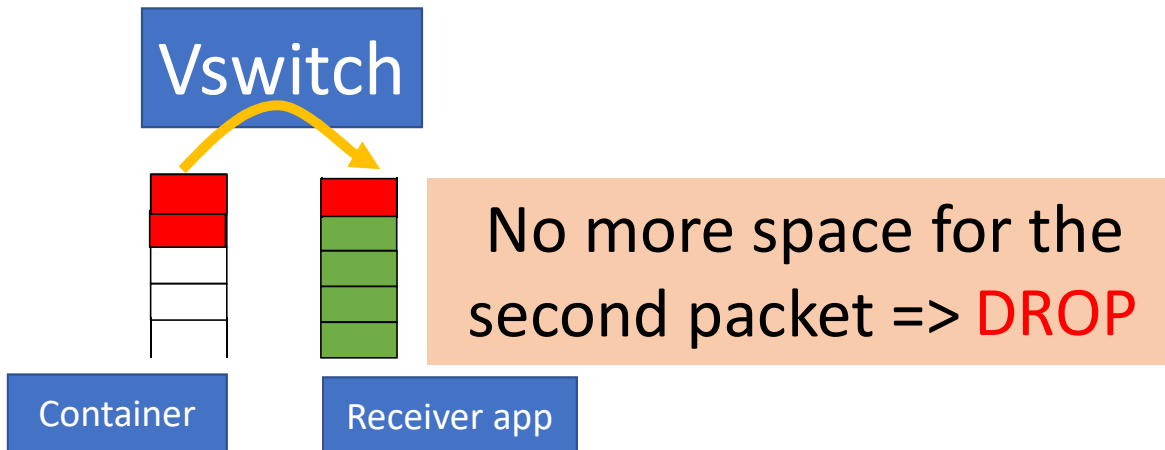
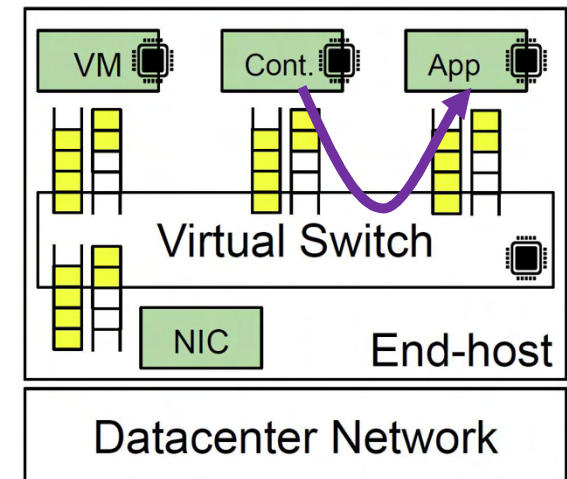
	Specialized Hardware	Performance Variability	Guaranteed Line Rate	Buffering Space	Queue Interface Principle
Physical switches	Yes	No	Yes	Limited	Push-based
Virtual switches	No	Yes	No	Unlimited	Pull-based

Push-based → In **physical switches**, the **switch itself** pushes packets from the buffer towards the destination

Pull-based → In **virtual switches**, **applications** pull packets from the allocated queues

*In virtual switches, it is the application that must pull packets from the queues **independent** of the vswitch*

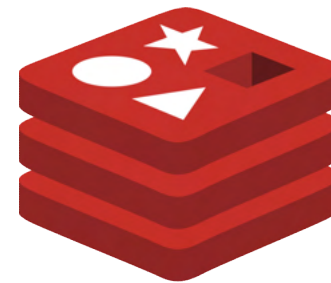
Packet loss at the end-hosts



Why is DROPPING bad?
Retransmission which includes buffer allocation, serialization, encapsulation, transmission.

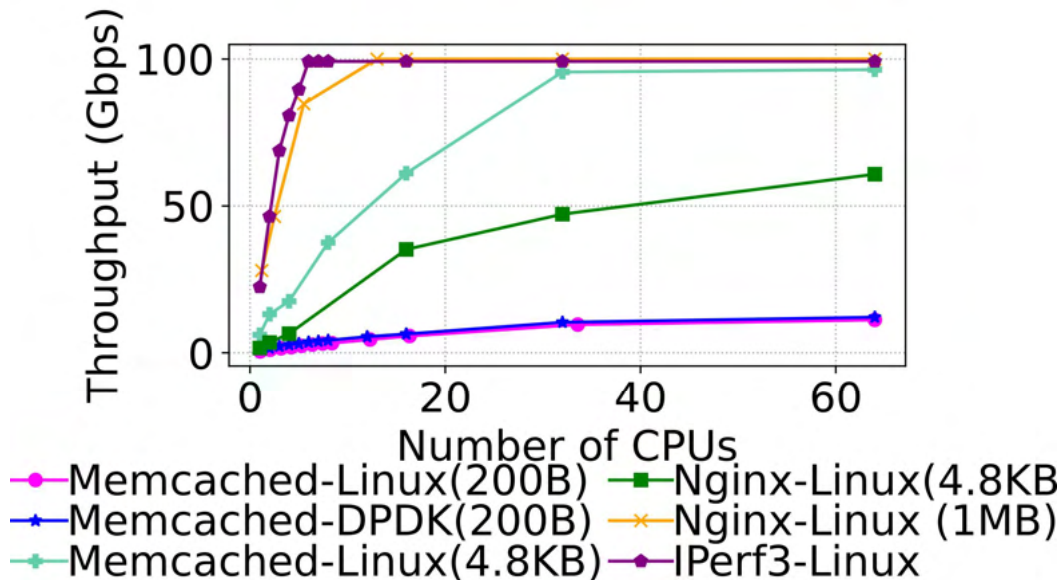
Slow receiver problem

- What is SLOW RECEIVER PROBLEM?
 - When an application is not able to keep up with the incoming rate
- What kind of applications are potentially slow receivers?
 - Any applications can be a slow receiver

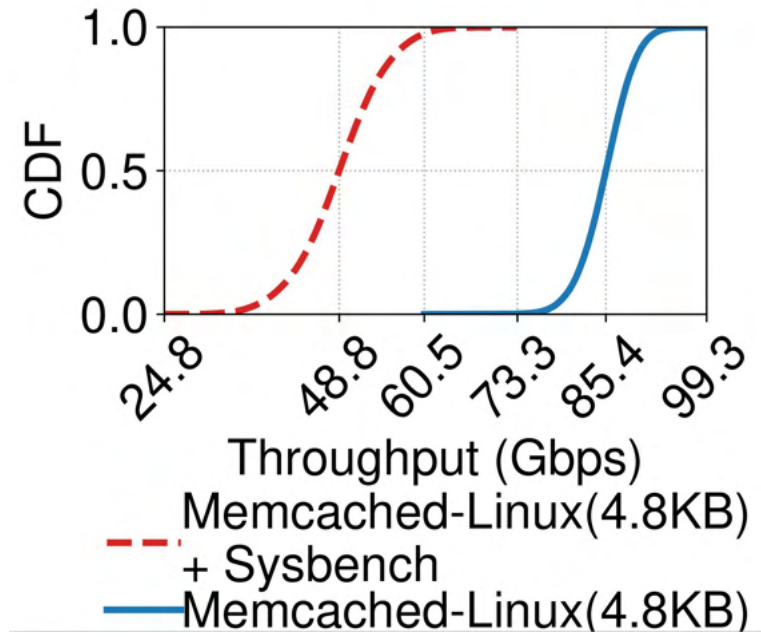


What causes Slow Receivers?

Single core application design

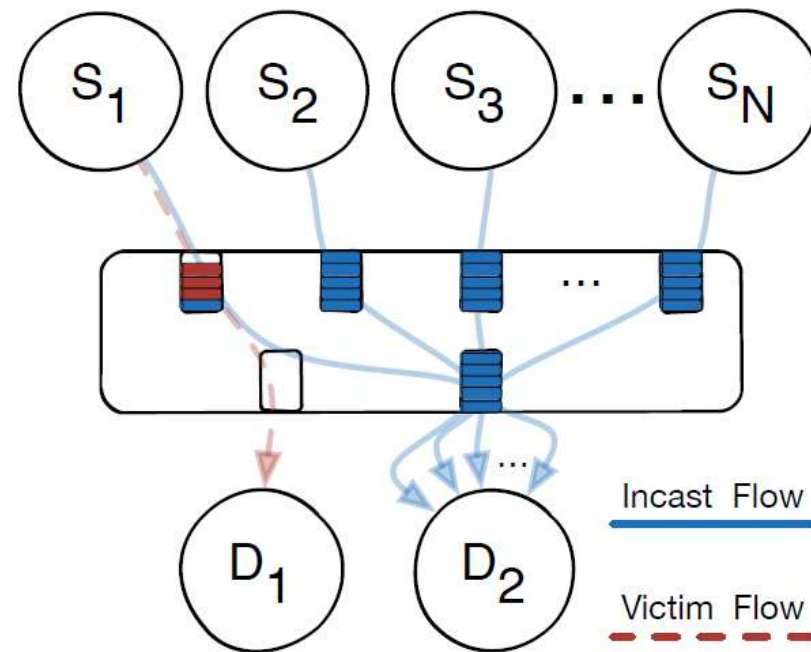
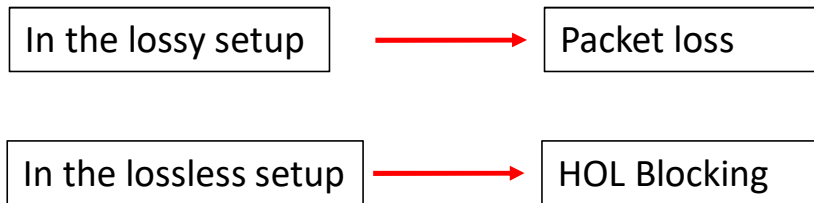


Application collocation

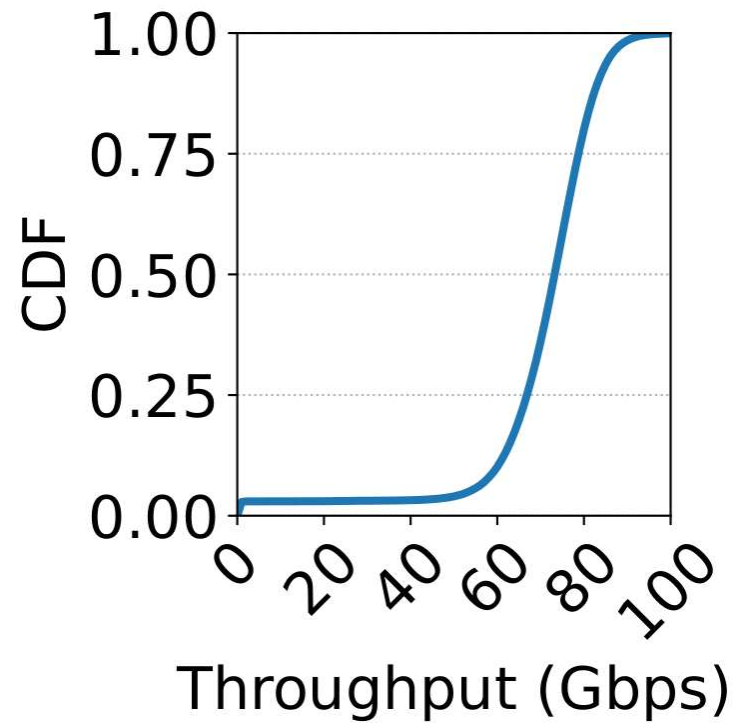
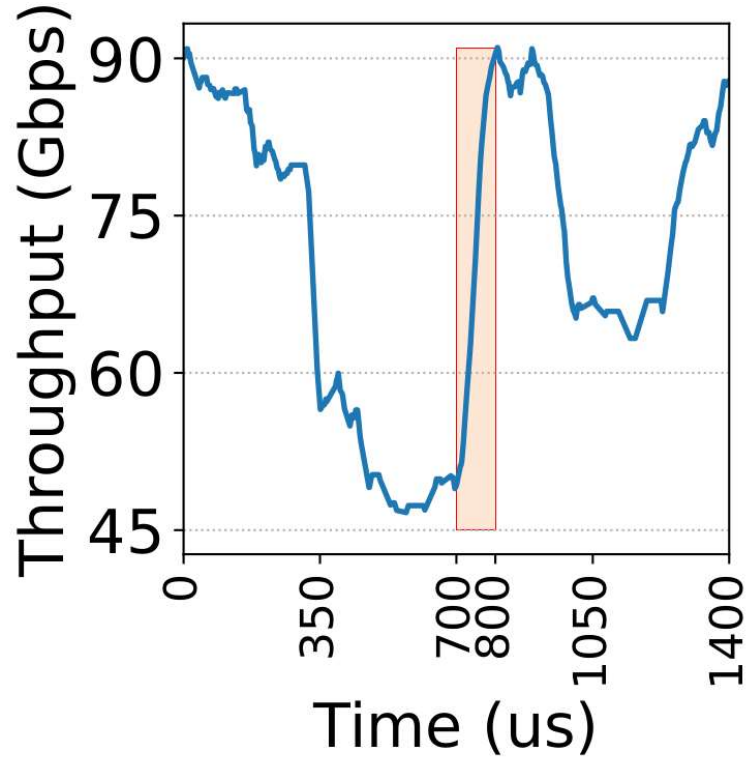


One core is simply not enough to fill 100Gbps.

Slow receivers cause either **HOL blocking** or **packet loss**



Can congestion controls address slow receiver problem?

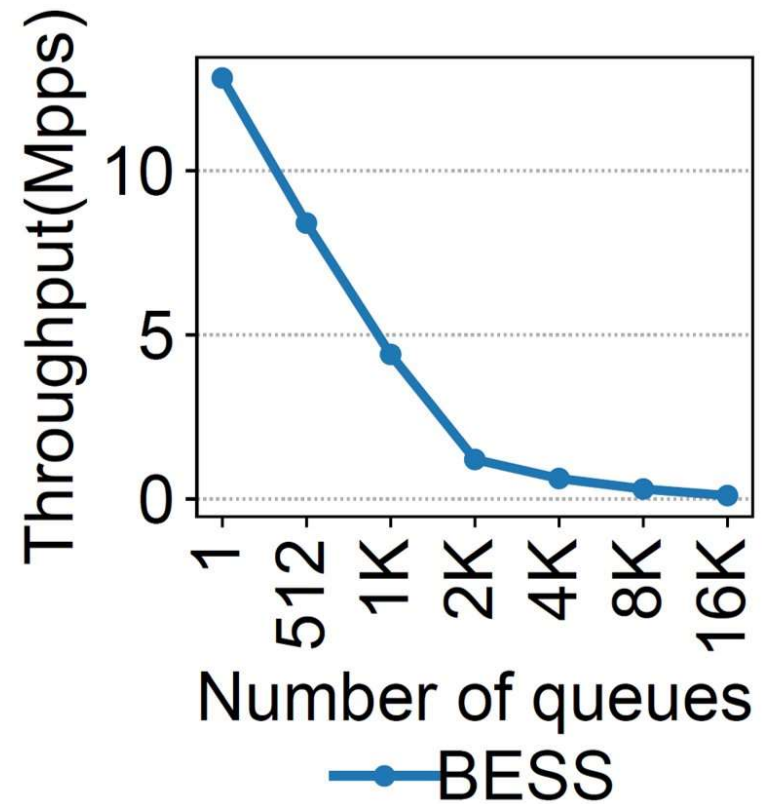
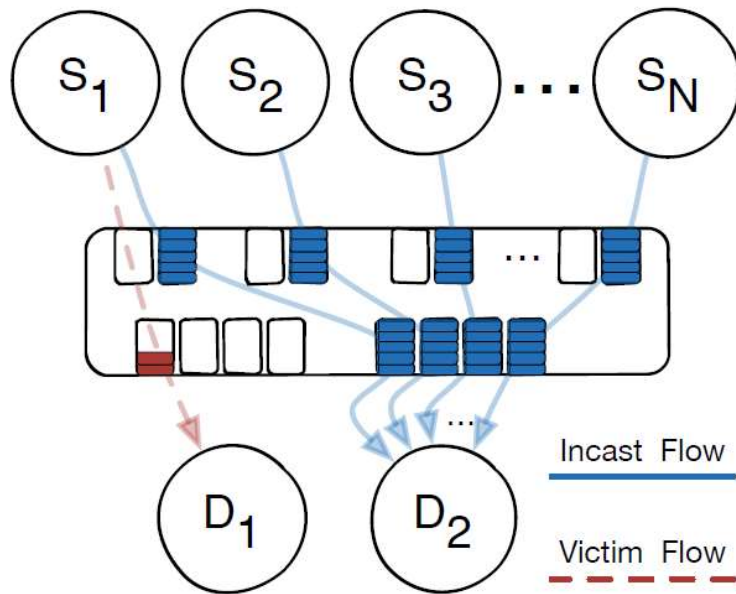


Variability at **SubRTT** time spans.

Per flow queuing

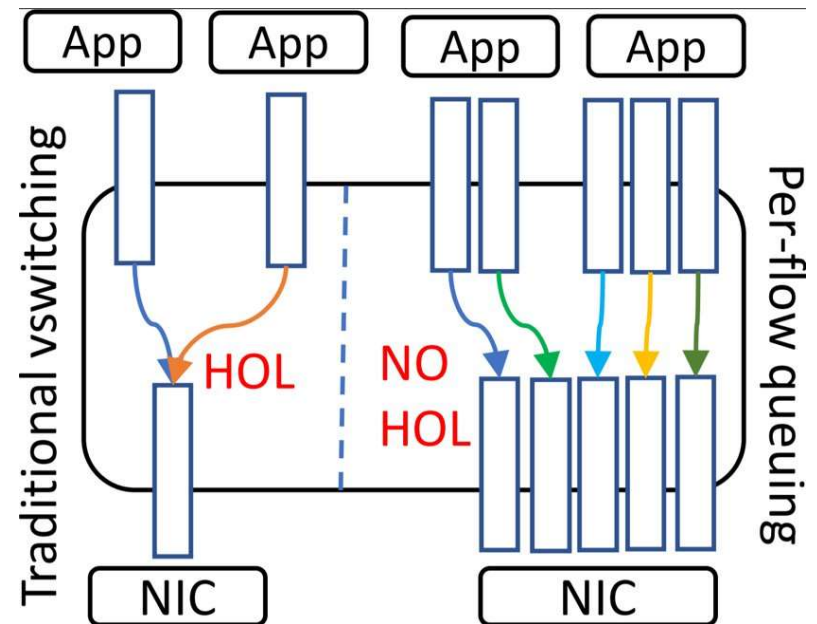
Is that it?

There is a challenge: CPU (polling overhead)



Where is this overhead coming from?

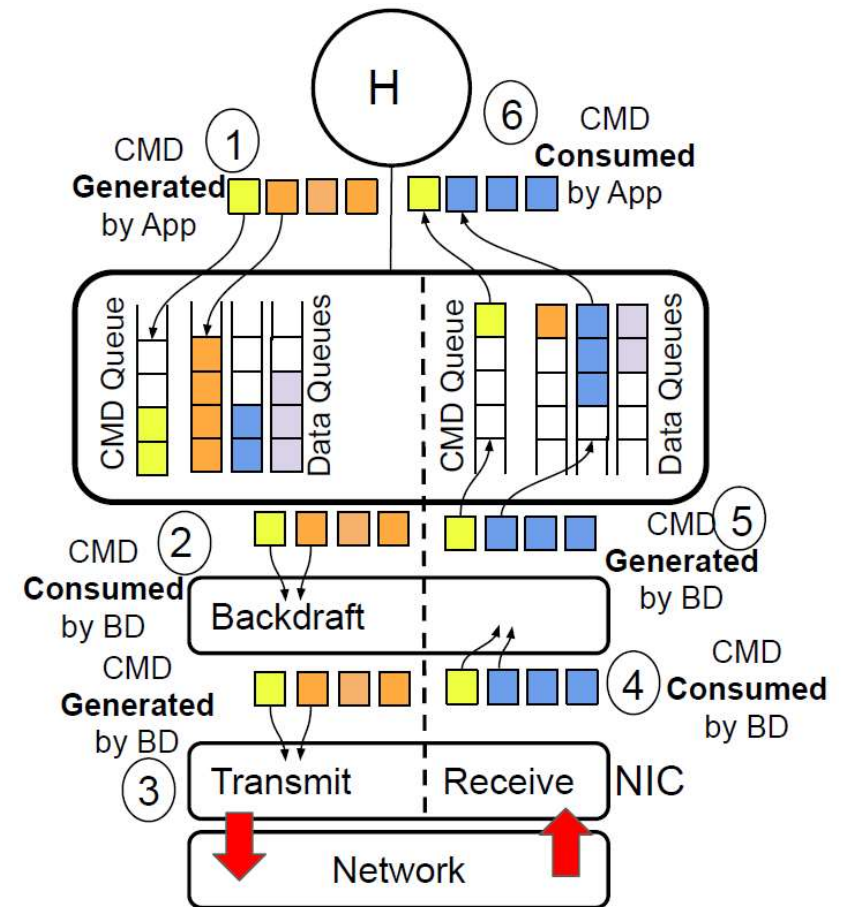
	HOL Blocking	High CPU Overhead
Large number of queues	No	Yes
Small number of queues	Yes	No



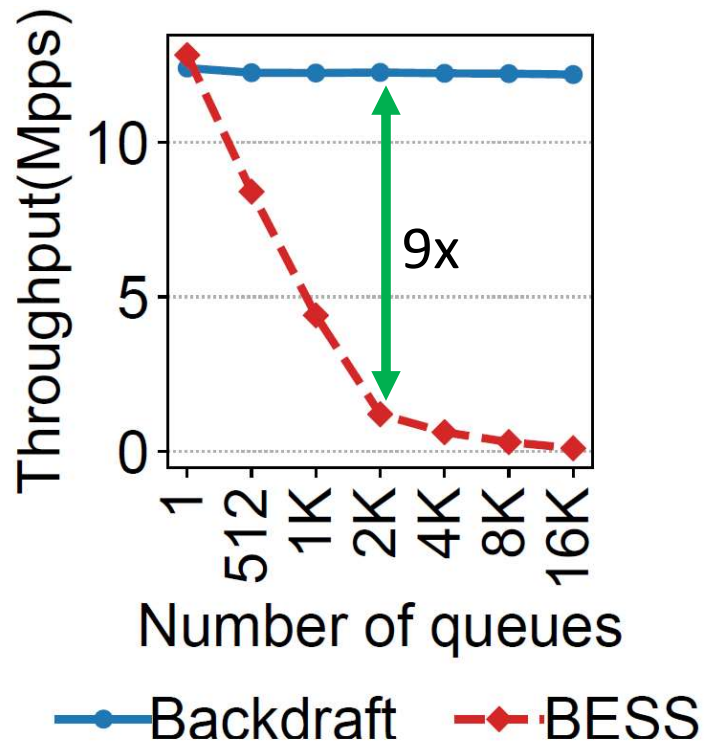
Doorbell queue

- CPU utilization optimization
 - Saves CPU cycles
 - Saves network bandwidth
 - Alleviates the slow receiver problem

The idea is to poll **only one queue** instead of so many.



Preliminary Results



We built Backdraft with both components on BESS:

- Per Flow Queuing
- Doorbell Queues

Question?