ERA: ECN-Ratio-Based Congestion Control in Datacenter Networks

Yukun Zhou[1], Dezun Dong[1], Zhengbin Pang[1], Junhong Ye[2], Feng Jin[2]
[1]National University of Defense Technology & [2]Tencent Inc

Background

- The widespread deployment of RDMA increases the stringency for convergence speed when congestion occurs in DCNs.
- Shorter propagation delay makes the queueing delay a major part of end-to-end latency in DCNs.
- Fast convergence and low buffer occupancy become more essential for lowering FCTs.

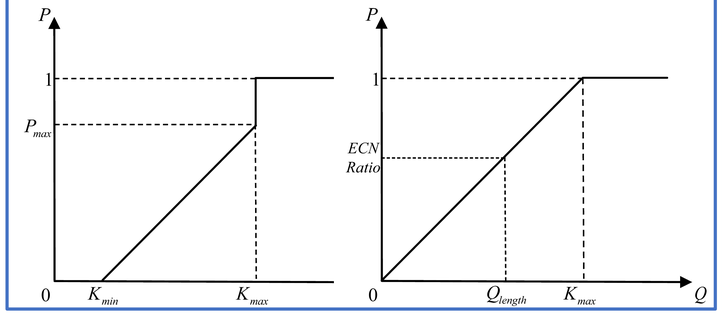
Introduction

We decompose ERA into three parts: the sender (RP), the switch (CP), and the receiver, (NP).

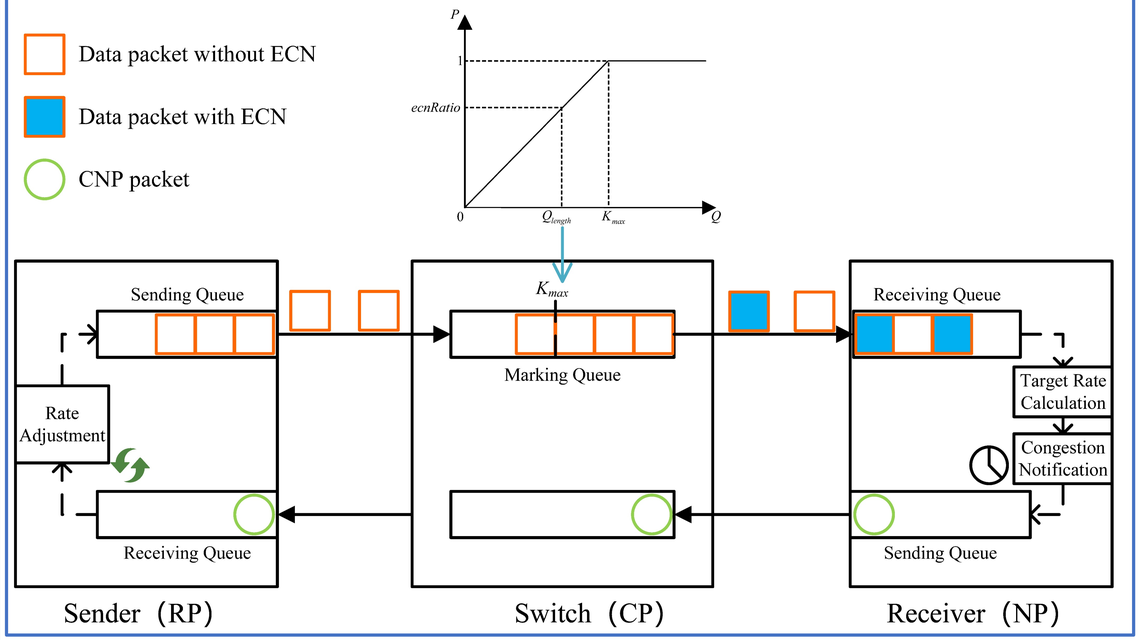
- The CP will mark packets with ECN selectively.
- The NP will record the ratio of packets marked with ECN and calculate the queue building-up rate and receiving rate so as to get the target rate for rate adjustment.
- The RP will adjust the sending rate according to the information in CNPs.

QBOE Solution

$$QueueLength = ecnRatio * K_{max}$$

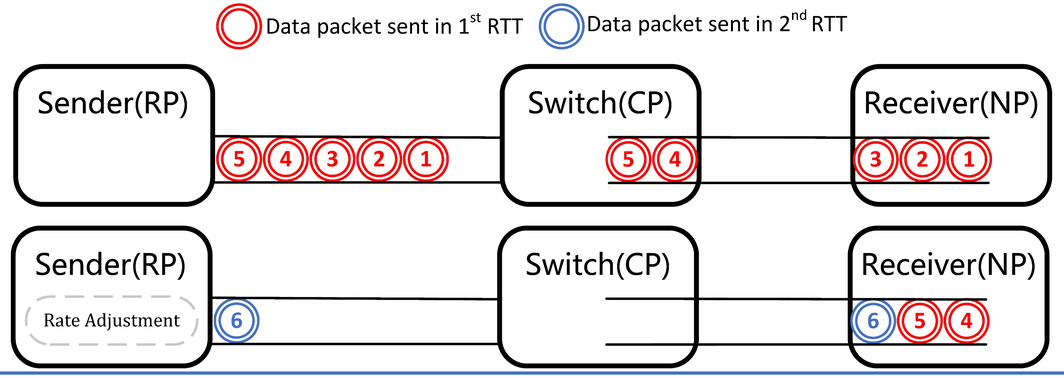


ERA Design



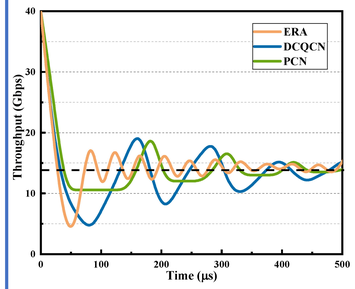
QDRA mechanism

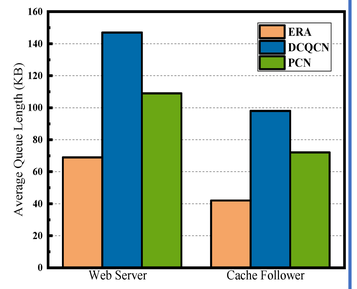
$$targetRate = receivingRate - queueBuildingRate$$

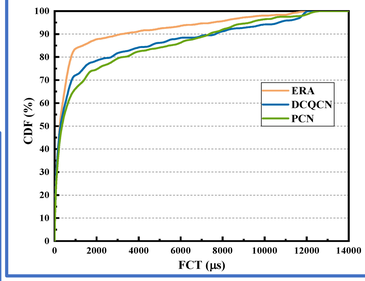


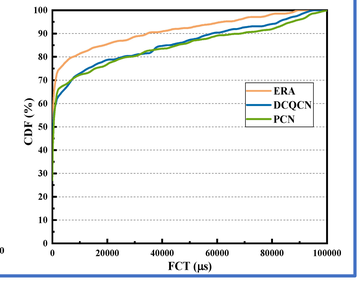
Evaluation

We use the OMNeT++ event simulator with INET framework to evaluate ERA in a variety of settings and compare it against DCQCN and PCN.
We measure the performance from convergence speed, flow complete time, and buffer occupancy.









Conclusion

The root of fast convergence lies in emptying the queue buffer occupancy generated in the first RTT in the minimum control periods in DCNS.
ERA outperforms DCQCN and PCN in convergence speed, flow complete time, and buffer occupancy with neither extra overhead for packet nor functionality addition in switches.