13 REXFORD

EScmp-Flowlet

We extend ECMP Flowlets by also using paths that have a Similar cost to the Shortest Path.

- Distribute the load over more links
- Problem: Possibility of cycles & increased delay
- Solution: Only allow a limited number of SCMP splits

Queue Length Estimator(s)

We estimate the queue length behind the switch to implement Random Early Detection. This is done by taking the max of two Queue Length Estimators:

- Counter Based: Counter in the egress to continuously estimate the length in the controller
- Meter Based: Use several meters with burst sizes equivalent to different queue lengths to estimate a lower bound

TCP global synchronization

We protect from global TCP synchronization in a congested network and prevent overly large TCP burst sizes

- Drop one TCP packet per flow every 0.1s once 80% threshold of bandwidth usage is reached

Lazy Heartbeats and FRR

How to efficiently detect links failure and recovery, without flooding the network?

- We use regular traffic to deliver “heartbeat” information: only send heartbeats lazily.
- Around 34% fewer heartbeat messages exchanged.
- Upon failure, fast re-route packets on LFA or RLFA.

Testing

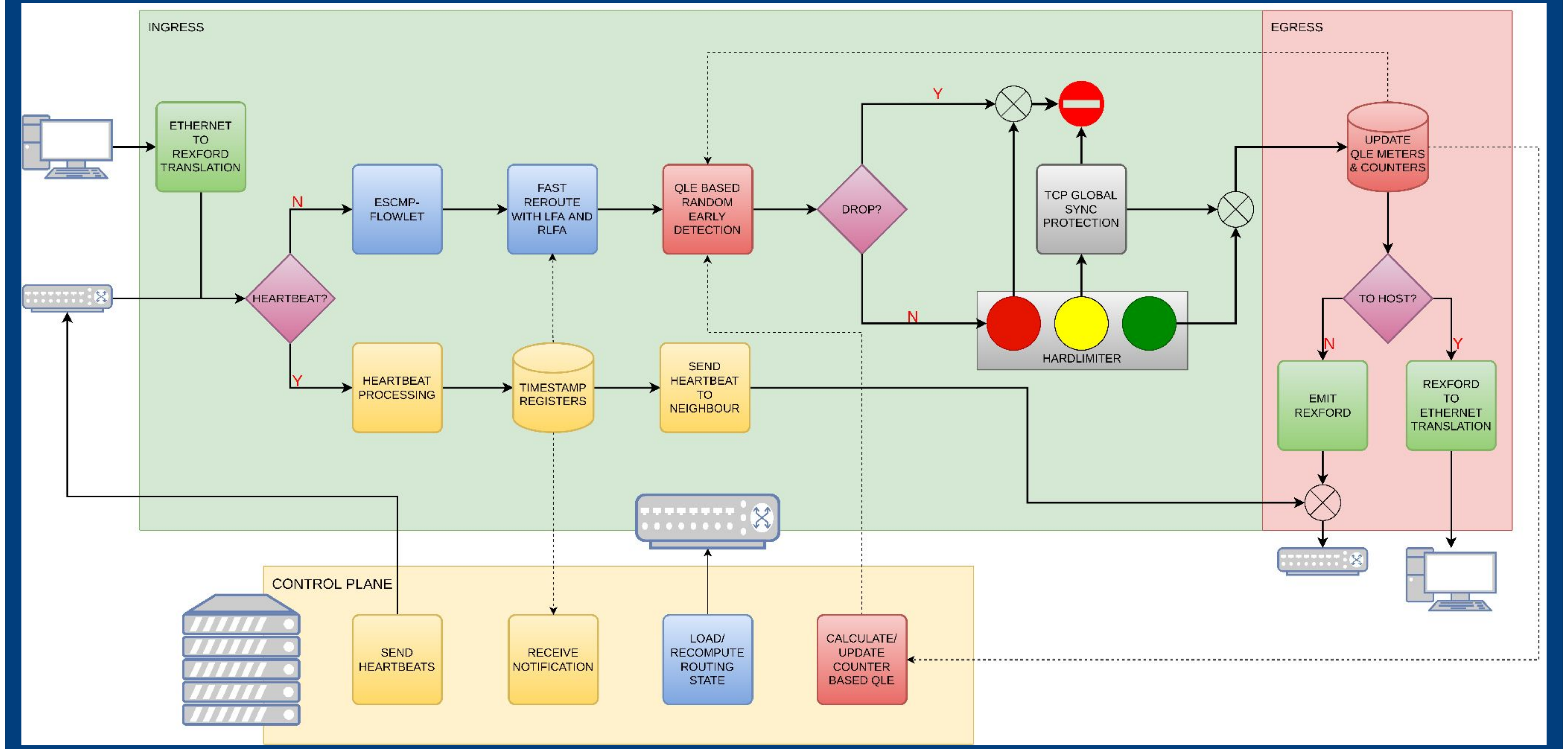
The tests were run with test.traffic as baseline, our group additional links and no failures.

- Heartbeats are the ones emitted at egress and logged by the switches.
- UDP delays represent the delay of the UDP flows as reported in results.csv. Same applies for TCP flow completion time.
- The measurements for UDP and TCP were repeated 5 times.

Using **Minimum Hop-Routing** to outperform minimum delay path routing.

Using **Queue length Estimators** to implement Random Early Detection.

Using **Lazy Heartbeats** for failure detection and fast recovery with minimal network overhead.



Bungeroth Matthias, Intoci Francesco, Westermann Floris

Results