Mitigating Congestions in Large-Scale LLM Training with RoCEv2 Feedback and SRv6

Arianna Quinci[1], Angelo Tulumello[2], Andrea Mayer[1,4], Paolo Lungaroni[1], Giacomo Belocchi[1], Pablo Camarillo[3], Ahmed Abdelsalam[3], Clarence Filsfils[3], Stefano Salsano[1,2]

(1) University of Rome Tor Vergata, (2) CNIT, (3) Cisco Systems, (4) COMMON NET

1 Networks for AI - Challenges

- AI workloads need fast, efficient data exchange across data center nodes
- RDMA (e.g., RoCEv2) enables low-latency transfers by bypassing the CPU
- RoCEv2 assumes a lossless network → lacks built-in congestion control
- Congestion Control Limitations
  - PFC (Pause frames)
    → causes Head-of-Line blocking, spreading congestion.
  - DCQCN (ECN + PFC)
    → requires fine tuning; sensitive to bursty LLM traffic.
  - Advanced solutions (Timely, HPCC, Swift):
    → Use telemetry (e.g., INT, RTT, ACKs)
    → Require programmability in both **network** and **end-hosts**

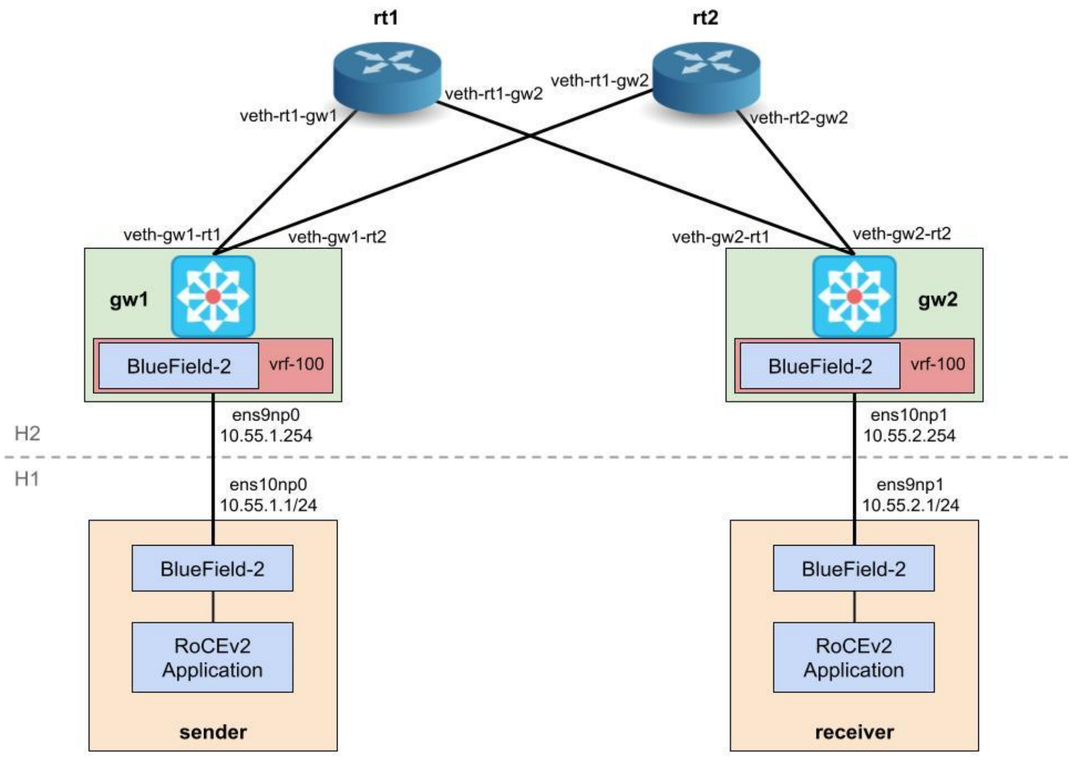
2 The role of programmability

In the network	In the end host
SRv6 fabric Datacenter networks traffic engineering & telemetry	**eBPF** for flexible packet processing implement/augment congestion control logic

3 Our approach

- **Goal:** Resolve congestion before performance suffers
- **How?**
  - (Near)Real-time congestion detection/avoidance
  - Use information from the network (telemetry, CNPs, RTT, ACKs...) to inform routing decisions
  - Reroute RoCEv2 traffic onto alternate ECMP paths
  - Using SRv6 to encode paths in the packet headers

4 Testbed



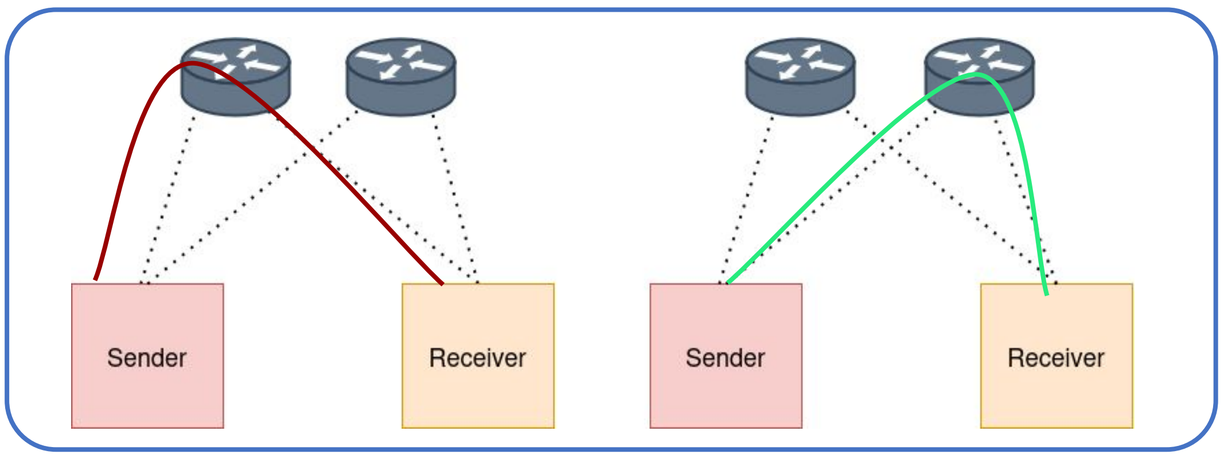
5 Experiments

Baseline Experiment

- Start RDMA on a chosen ECMP path
- Achieve **target rate** under normal conditions
- Continuous throughput monitoring

Adaptive rerouting under congestion

- Introduce manual congestion by throttling a virtual link
- Observe throughput drop
- Detect congestion & update SRv6 encap SID list

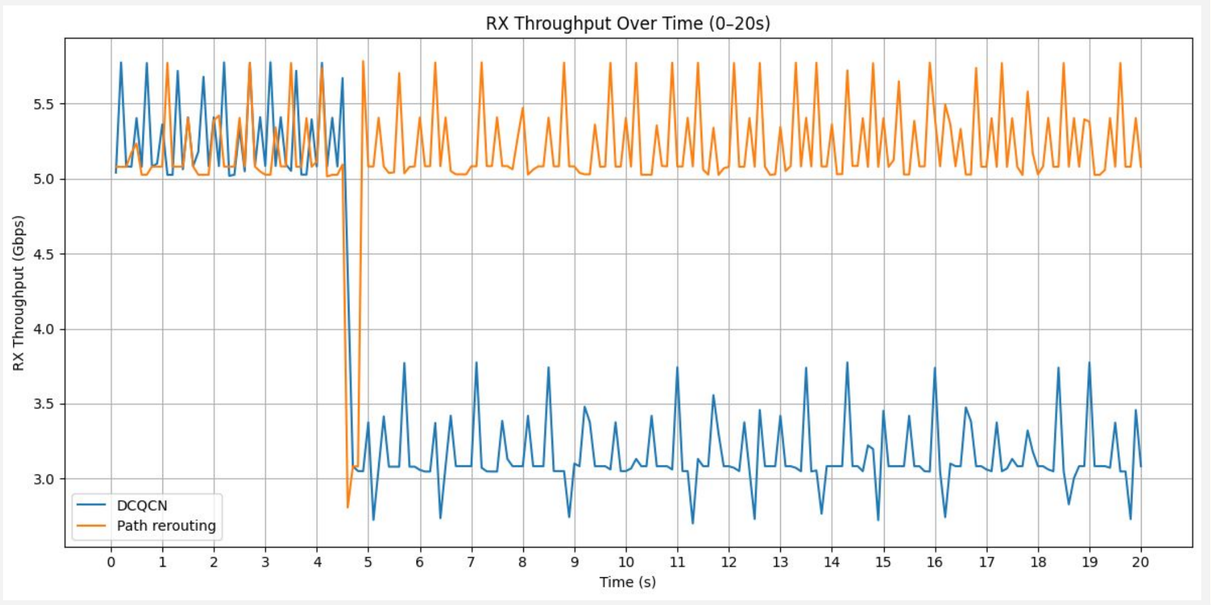


6 Implementation

- Our PoC uses eBPF/XDP on the host to:
  - SRv6 encap/decap
  - Copy congestion bits between inner/outer headers
  - Monitor RoCEv2 CNPs
- Adapt traffic flow based on congestion signals
  - Augment the congestion detection with both CC and network data

7 Results

- Monitoring on GW1 interface
- Sampling every 100ms
- Artificially throttling the link



ACKNOWLEDGEMENT
This work has received funding from the Cisco University Research Program Fund and from the EU under the Italian NRRP program “RESTART”.